

Failure need not mean despair

The collapse of the Reykjavik pre-summit at the weekend is mystifying, explicable only if both sides stood unreasonably on principles that could not have mattered.

REYKJAVIK seems not to be a happy place for encounters between Soviet and United States citizens. The splendid chess championship won last week in Moscow by Kasparov is a reminder of the earlier championships, more than a decade ago, when the then US champion, Bobby Fischer, won the match but lost the war in his contest with an earlier champion, Boris Spassky; Fischer's victory cost him his reputation as a serious person. Similarly, there were no winners, only losers, at Reykjavik last weekend. A settlement, or even the promise of a settlement, would have helped Mr Reagan's party's candidates at next month's mid-term election; presumably the failure will now work the other way, but not seriously. Similarly, for Mr Gorbachev, an agreement would have allowed faster progress with the overdue reform of the Soviet economy (and would have reduced the influence of the Soviet military on matters they know very little about). Failure will presumably at best postpone that happy outcome. Both men at Reykjavik will also have lost something in reputation, Mr Gorbachev with those of his colleagues in Moscow who have always been sceptical of his strategy, Mr Reagan with a wider group of well-wishers.

After the summit meeting that never was (Reykjavik was a warm-up only), it is important that the failure should not seem bigger than it is. So first count Reykjavik's pluses. The two sides are obviously within an ace of agreement on two important issues: long-range or strategic nuclear weapons (only half as many as at present) and missiles of intermediate range (a total of 100 on each side, none based in Europe). It should be a great encouragement that the two sides came close to agreement on the outlines of treaties regulating the two missile issues, although that is not the same as putting ink on paper.

On a third issue, the comprehensive test-ban that the Soviet Union is demanding, there should have been more room for compromise than was used at Reykjavik. Last Friday, the United States said that it is now prepared to ratify the two partial restraints on testing (the threshold test-ban and the regulation of

peaceful nuclear explosions); this it should have done at least six years ago, but the United States also stands by the belief that continuing tests are necessary to verify the utility of the stockpiles and to fit new rockets with new warheads. That argument is thin. Who, for example, would estimate the shelf-life of some other complicated machine in storage by testing single items every now and again? Statisticians would be appalled at the idea. Yet it would have cost very little to substitute for the concept of a comprehensive test-ban one of the variants of all-out restraint that have been current in the past few months, a small quota of tests every other year, for example. The United States, which was ready to sign a comprehensive test-ban treaty towards the lame-duck end of Mr Jimmy Carter's presidency, has backed away from that enthusiasm because of last weekend's real sticking point, the Strategic Defense Initiative (SDI).

Too little attention has been paid to that can of worms in the past few months. Two of the three sub-groups at the bilateral arms control negotiations at Geneva have, for the past few months, been purring like cats within sight of cream. It now emerges that the third sub-group has seen the ructions happen. Mr Gorbachev's demand at Reykjavik was for a technical amendment of the Anti-Ballistic Missile Treaty (signed only in 1978) that would clarify (and tighten) the definition of "research". Mr Reagan's offer at the United Nations last month (see *Nature* 323, 381; 1986) that SDI would not be deployed before 1992, and not then without two years of negotiation with the Soviet Union, was clearly part of the argument behind the scenes. Mr Gorbachev had earlier, perhaps misleadingly, said that "laboratory research" would be allowable, but without explaining whether his expanded definition would permit, say, the testing of X-ray lasers driven by nuclear explosions in deep holes underground. Last weekend, he was more explicit and too restrictive for Mr Reagan.

With all the muddle about SDI, we should all be clear why it is such a contentious issue. That the project will probably never succeed technically in the sense of which President Reagan appears to dream is irrelevant. General James Abrahamson, the director of the project, is fond of showing sceptical visitors a crude newspaper cartoon in which one puzzled US citizen asks another, "If it won't work, why are the Soviets so afraid of it?". The truth is that, while the system as a whole will never be as sure a defence against hostile missiles or, untested, as comforting as Mr Reagan hopes, some bits and pieces of it are likely to have considerable military value. At this stage, nobody can tell what these will be, but they are almost certain to give the first in the field (the United States) a military advantage. For example, one obvious product of the money now being spent is a better early-warning system based on infrared detection, doubling (from 20 to 40 minutes) the likely warning of the arrival of hostile nuclear weapons, and lending credence to the hair-trigger strategy of "launch under attack" (the concept of retaliation before being hit). Another is a better system of terminal defence against missiles than the US Sprint system abandoned in the early 1970s.

But even Mr Gorbachev would have to admit that, without a means of destroying Soviet missiles *en route*, such a system

News: Contents

US nuclear safety: Hanford plutonium plants in a narrow squeak	569
Pharmacia takes over LKB Producter	
US biotechnology: Congress at odds over regulation	570
Soviet education: Central role for the universities	
International space: Space year for Columbus	571
Japan fellows: British are reluctant travellers	
UK fund for China	
Direct broadcasting: Unravelling an encryption code	572
Nobel prizes: Growth factors bring rewards	
Asylum: US cancer researcher in Moscow	573
"Extinct" lemurs found in Madagascar	
British research: Better measures of decline	
European research: Joint centres in eye of storm	574
Japanese astronomy: Projects wait on organization	
US astronomy: Big ideas, short pockets abound	575
French research: Optimists and pessimists	



would not allow the United States to launch what the Soviet Union fears most, a first strike against military targets. That is why it matters that SDI will probably also give the United States some facility for destroying Soviet missiles in flight, most probably not by fancy laser technology but by more old-fashioned ways of destroying unwanted objects. Mr Reagan seems unreasonably insensitive to Soviet suspicions that, in this guise, SDI would be a convenient umbrella for a first strike.

This, a constant theme at the negotiations at Geneva, and probably also at Reykjavik, has been a serious issue from the outset. Mrs Margaret Thatcher, the British Prime Minister, went to Washington three years ago to secure a promise that SDI is, for the time being, a research programme only; then, as now, the United States seemed not properly to appreciate that, without such an understanding, SDI research could do endless damage. Mrs Thatcher's understanding of "research" may have been different from Mr Gorbachev's, but they have been making the same point. (If the United States concedes that the Soviet Union is entitled to some assurance on the point, should it not also concede the need for verification of SDI, one of the constant complaints at Soviet proposals for arms control?) But Mr Reagan's statement to the United Nations last month (see *Nature* 323, 381; 1986) was a more explicit account than previously of the obligations the United States would assume.

So what went wrong at Reykjavik? On the face of things, the gap between the two sides could have been bridged by negotiation. The objective should have been to find a way of containing the SDI programme within reasonable limits for the time being, with the understanding that the future of the project would be decided at some pre-set date. This position, not very different from what Mr Reagan offered at the United Nations or from what Mr Gorbachev was asking at the weekend, would have allowed Mr Reagan to continue making speeches about the need to abolish the threat of nuclear weapons and Mr Gorbachev to keep on saying that SDI is an abomination. History will tell whose small point of principle proved insurmountable at Reykjavik, and whether it was worth the trouble it will cause.

What will that be? The immediate danger is a shouting-match, a return to the early 1980s. Both sides have learned too much for that to happen. But the abandonment of the Geneva talks, which would not surprise despondent post-Reykjavik George Shultz, the US Secretary of State, would be a bad sign. The middle-term danger is what it has always been, that relationships between the two superpowers will deteriorate to the point at which war in Europe, with all its new-fangled engines, comes to seem again an ever-present threat. But is that not the case now? The oddest feature of the past few years is that, while the names that the two sides have called each other have been no more tolerable than in previous decades, war in Europe has seemed less imminent. Why should that be? Because the Soviet government, whatever its intentions, has found it necessary to relax its supposition that the world is divided into two by the Elbe.

People on both sides of the mutual German frontier now have a more vivid understanding of what the others have to offer. In the technical community, the continuing sad isolation of Soviet science is nothing less than tragic — for both sides. A week or so after the release of Orlov from prison (and his unkind expulsion from the Soviet Union), it is worth remembering that even arms control is merely a means to ends still undefined. □

Underpaid and overtaxed

The 1986 Tax Reform Act will make the United States even more a magnet to technical people.

A FEW days after President Ronald Reagan signed into law the 1986 Tax Reform Act, the British Chancellor of the Exchequer, Mr Nigel Lawson, was last week telling the annual conference of the Conservative (and government) Party that he still intends to see the basic rate of income tax, at present 29 per cent, reduced

to 25 per cent. Under the British system, the basic rate is that which affects or afflicts most taxpayers; those who earn substantially more than the average salary pay at a higher rate (at the margin) up to a maximum of 60 per cent. Mr Lawson's declaration appears to have been ecstatically received, not least as a sign that the British government is sticking by its seven-year-old election promise. But neither he nor his audience appears to have noticed that the effect of the US Tax Reform Act is to reduce the highest rate of personal taxation in the United States to below what the British chancellor must for the time being call "standard". Nothing could more starkly illustrate the differences between these two English-speaking countries, or the reasons why one of them is economically and even technically more successful than the other.

To be fair, the Tax Reform Act is not the unalloyed blessing it may seem from a distance. Many of the tax shelters in which the middle classes of the United States have tucked their savings will in future be taxable, while there will also be limits on the amounts of interest paid (except on mortgages on first and second homes) that may be set off against income before the computation of tax. Medical and educational spending (above certain limits) will not in future be wholly a charge on the US Internal Revenue Service (better known as IRS), while there will be extra costs to pay for state and local taxes. The restraints on what counts as a charitable deduction from income will hurt some charitable institutions, universities among them, while there is nothing in the tax bill to help reduce the budget deficit (which would require an upper rate more like 30 than 28 per cent). The interim plan that corporations should pay taxes at a higher rate (32 per cent) could damage economic growth, but there is an army of accountants and corporate treasurers seeking ways round that obstacle. What the reform act will do is to confirm the general impression of what is the reality, that the United States is a low-tax society. That will further exacerbate a string of problems with which the high-tax societies of Western Europe have been living for several years.

The chief of these is the readiness of technical people to move from one place to another. For the past two decades, successive British governments have been wringing their hands over the rate at which technically trained people have departed for the United States, usually with seemly protestations that they are not going for the money but for the professional opportunities but in the certain knowledge that their skills will also command a higher price. Prudent would-be emigrants from Western Europe, to be sure, have always known that the balance of advantage is not calculable by simple arithmetic; even in Britain, the value of the social benefits provided by what the American Medical Association calls "socialized medicine", for example, are far from negligible. But the Tax Reform Act cannot but accelerate the emigration, at least while tighter US immigration laws seem not seriously to have restricted the freedom of US organizations to hire skills in short supply (nor should they).

How will Mr Lawson and his government colleagues respond to that prospect? The obvious danger is that they will think it unimportant. They may also, plausibly, complain that it would be ridiculous, and economically impractical in any case, for the government of Britain or some other Western European state to amend its own taxation policy so as to compete more effectively with the United States for the allegiance of a small proportion of its population. But that, unfortunately, is only half the story. With time, the market in technically qualified labour is bound to become more international and more free. In the long run, places that underpay and overtax their technical people will find themselves left only with those who lack the qualifications to get up and go, hardly the best guarantee of technical or economic success. So, in the long run, there is a sense in which Mr Lawson (or his successors) will have no choice but to emulate the US 1986 Tax Reform Act. That is yet another reason for believing that this piece of legislation may be the most enduring and influential monument to the Reagan administration. □

US nuclear safety

Hanford plutonium plants in a narrow squeak

Washington

PUBLIC apprehension about nuclear safety was stirred again last week when the Department of Energy (DoE) shut down two military plutonium plants at the Hanford nuclear reservation in the state of Washington. Operations will be halted until the outcome of an investigation into widespread safety violations which was set up after an incident on 29 September in which rules designed to avoid the formation of a critical mass of fissile material were broken.

DoE's order to stop operations followed only by a few days the publication in the *Seattle Times* of highly critical internal safety audits of the Hanford plants conducted by the operating contractor, Rockwell-Hanford operations, a division of Rockwell International. DoE denies any link between the leaking of the audits and the decision to shut down the two plants, which could remain out of action for a month or more. Other nuclear facilities at Hanford run by other contractors are not affected.

The two plants concerned are the PUREX plant, which chemically separates uranium and plutonium from irradiated nuclear fuel, and the plutonium metal finishing or "Z" plant, which converts plutonium nitrate solution from the PUREX plant into metallic plutonium for nuclear weapons. The irradiated fuel com-

es from the "N" reactor, also at Hanford, which was itself the subject of a recent DoE safety inquiry because of its similarities to the Soviet reactor at Chernobyl that exploded last April.

During the 1970s, the Rockwell-



Hanford's PUREX processing plant.

Hanford plants were producing only small amounts of plutonium for research, but production for weapons began again in earnest in the early 1980s. In the 29 September incident, workers in the "Z" plant started to transfer a plutonium-bearing solution through a temporary pipeline isolated by five manually operated valves from a tank which was not "critically safe" in the sense that it could not be filled with the solution without the risk of its becoming critical.

Regulations require fissile solutions to be isolated from unsafe tanks by fixed

plugs. Criticality would produce intense radiation and large amounts of heat, and could cause an explosion; at least five people have been killed in the United States in criticality accidents since the 1940s.

None of the solution actually reached the unsafe tank, but even so, Rockwell rated the incident at 4 on a 1 to 5 scale of seriousness. The incident was reported to DoE on 2 October, and the shutdown order came six days later, after DoE inquiries found other instances of by-passed safety controls and a general "lack of appropriate controls, reviews and procedures".

The audits leaked to the *Seattle Times* indicate a pattern of routine violations of safety measures and of the rules devised to guard against theft of nuclear material.

Violations of the redundant safety regulations to ensure against solutions going critical are not themselves rare; 23 were reported during the year that ended on 1 October, which led Representative Ron Wyden (Democrat, Oregon) to claim last week that it was publication of the critical safety audits and the prospect of congressional interest, rather than the 29 September incident, that led DoE to decide to put its house in order.

The auditor whose critical conclusions were leaked, Carl Ruud, spent last Thursday in closed discussions with the oversight and investigations subcommittee of the House of Representatives' Energy and Commerce Committee, which has been investigating safety at Hanford for some months. Ruud is likely to be invited back to give formal testimony. Wyden said the shutdown showed the seriousness of the safety problems at Hanford and that DoE's safety reviews had "lost all credibility". Ruud, who still works for Rockwell, has declined to speak to the press.

The Hanford reservation, which covers more than 500 square miles, has been in the plutonium business since the beginning; it produced the material for the Nagasaki bomb. And that may be its biggest problem: many of the plants are still using Second World War technology.

The site is also dirty, in both the conventional and nuclear senses. Documents released during the past year indicate that Hanford released up to 1 million curies of radioiodine and kilogram quantities of plutonium into the environment during the early years, causing "Chernobyl-sized" doses to milk-drinkers in the vicinity, according to Robert Alvarez of EPI. Hanford now has huge problems with ground contaminated with chemical solvents and accumulating high-level waste sitting as sludge in decaying storage tanks; one estimate puts the cost of cleaning up the site at not less than \$11,000 million.

With increasingly stringent oversight from Congress, DoE may have to face the indefinite closure of some of the plants.

Tim Beardsley

Pharmacia takes over LKB Producter

PHARMACIA AB, one of Sweden's leading pharmaceutical companies but best known to the laboratory scientist as the manufacturer of materials for molecular separation methods, is to buy up LKB Producter AB, a rival Swedish company, best known for its high quality instruments for separation techniques. The deal represents a considerable about-face for Pharmacia, which was itself the subject of a takeover bid last February, which was stopped only after Refaat El-Sayed, president of Fermenta AB, the company planning to buy Pharmacia, admitted falsifying his claims to a master's and a PhD degree.

Already the two leading European companies in laboratory separation, Pharmacia and LKB are together estimated to account for at least 40 per cent of a world market of US\$560 million a year. Electrophoresis and chromatographic techniques and equipment are the mainstays of their business. But both companies have recently developed interests in diagnostic

kits, with LKB taking a lead in the development of immunoassay kits that use non-radioactively labelled reagents.

Last week, Pharmacia agreed to acquire 82 per cent of LKB's voting stock from the holding company Incentive AB for 776 million Swedish kroner (\$113 million). It now plans to obtain the remaining shares for another \$73 million. In the past three months, Pharmacia has also acquired Intermedics Intraocular Inc., a Californian eye-care company, and AB Leo, a Swedish drug company that cost it \$480 million.

Meanwhile El-Sayed, who remains majority shareholder and group chief executive of Fermenta, last month signed agreements to sell a third of the voting shares in the company to three Swedish groups. The meaning of the sales in the run-up to the decision in November by Montedison, the Italian chemicals group, to whether to pursue its earlier plans to acquire a majority stake in Fermenta, is hard to fathom.

Peter Newmark

US biotechnology

Congress at odds over regulation

Washington

ELECTION fever has turned criticizing the US government's plans for regulating biotechnology into a political spectator sport in recent weeks, with points awarded for arcane theoretical arguments from ecology and molecular biology. But the game turned acrimonious last week, with each side accusing the other of bias, and worse, over a report on the subject prepared by staff of the investigations and oversight subcommittee of the House of Representatives' Committee on Science and Technology. The government's proposals, published in June, are designed to provide a framework within which federal agencies could coordinate their regulation of biotechnology — and which would let the industry know where it stands.

The report is based on hearings held by the subcommittee during the past four months. Considering that it has not been released to the public and that few admit knowing its precise contents, it has engendered strong feelings. But defenders of the administration's "coordinated framework" for regulation published in June have been given a clue to what to expect by comments on the administration's proposals published by Representative George Brown (Democrat, California) and four other Democratic members of the subcommittee, including its chairman Harold Volkmer.

The comments are stridently critical of the "coordinated framework" and have given rise to alarm that the formal report will have a "negative tone". One reason for that fear is that the report, although without direct legislative impact, might influence the forthcoming lawsuit against the administration brought by environmental activist Jeremy Rifkin. Rifkin claims that the "framework" does not meet the requirements of the Environmental Policy Act. The comments also contain errors of fact, according to Dr David Kingsbury, head of the Biotechnology Science Coordinating Committee, which has the job of implementing the "framework". Staff say, however, that there are important differences between the congressmen's comments and the forthcoming report.

Last week, the full Science and Technology Committee balked at giving its approval to the staff report, with minority Republicans saying the work was seriously flawed. An attempt to shelve the report until next year was defeated, however, and staff now expect approval, possibly with some minor changes, before Congress adjourns this week. Volkmer, who represents Missouri, supports the staff report. Observers point out that the Monsanto Company, which has received un-

favourable publicity over its proposals to field-test a genetically engineered microbial pesticide, is prominent in the state, and that memories of Times Beach, where industrial pollution forced a community to be evacuated, still linger.

One particular scientific issue seems to have divided Democrat from Republican on the subcommittee: whether recombinant microorganisms whose new DNA includes only non-coding regulatory sequences should be exempt from high-level safety review before release into the environment. The administration, in the "framework", says they should. But in formal comments to the White House, Monica Riley of the American Society for Microbiology (ASM) said that there are known cases of foreign regulatory sequences affecting an organism's host range, and that they can affect specificity. The society would not choose automatically to exclude such recombinant organisms from high-level review. Other criticisms relate to exemptions for opportunistic pathogens.

Riley's comments, and testimony in Congress have earned ASM some criticism. In a letter addressed to ASM president Dr Jean Brenchley and dated 3 September, Winston Brill, vice-president for research at Agracetus Inc., and Ronald Cape, chairman of Cetus Corporation, say the ASM position is not supported by "the knowledgeable majority" and that "the portions of the ASM testimony [to the subcommittee] that are incorrect should be corrected by ASM; however, it may be impossible to retrieve the situation". Harlyn Halvorson of Brandeis University, chairman of the relevant ASM committee, said a reply to Brill and Cape's "arrogant" letter was being drafted but that the society was unlikely to change its position.

Meanwhile, the Office of Science and Technology Policy is weighing an 18-inch stack of comments it has received on the "framework", which outlines possible regulatory roles of all federal agencies involved in biotechnology (see *Nature* 323, 387; 1986). Although most of the comments are less harsh than those of Representative Brown and his colleagues, the complex proposals have given rise to numerous criticisms. One possibility, prompted in part by researchers' concern about the different sets of research guidelines envisaged in the "framework", is a uniform agency-wide set of recombinant DNA research guidelines that would encompass both those of the National Institutes of Health and proposals put forward by the Department of Agriculture. That proposal will shortly be considered by the Biotechnology Science Coordinating Committee.

Tim Beardsley

Soviet education

Central role for the universities

THE Soviet minister of higher and specialized secondary education is "completely dissatisfied" with the research output of the higher education sector. Minister Gennadii Yagodin says that although the higher education sector employs some 35 per cent of the country's scientists (including about half of all the doctors of science), it is responsible for only 10 per cent of the total Soviet research effort. This, the minister admitted, is not entirely the fault of the universities, "but also their plight", because in recent years serious deficiencies have been allowed to develop in the "material and technical base" (buildings and equipment) of the higher education sector. Nevertheless, whoever is at fault, the fact remains that the university research sector is failing to "justify itself within the state".

The Politburo of the Central Committee of the Communist Party recently issued a set of guidelines for the restructuring of higher education, Yagodin said, and during the coming academic year, the universities and higher educational establishments will have to take the first step towards this "restructuring".

But at the planning level, the first step seems to be at a standstill. For the first time for many years, the total number of universities and institutes of higher education in the Soviet Union is being held steady, at 894, with a total student body of around 5 million. This, Yagodin explained, marks a shift in emphasis from quantity to quality. "Negative qualities", he said, have developed in Soviet higher education in recent years because of an emphasis on "quantitative indicators".

At the research level — the target of Yagodin's (and the authorities') particular disapproval — research topics are to be "revised" to make them more useful for practical purposes, and undergraduates are to be encouraged to participate in their instructors' research projects.

Yagodin's remarks seem to reveal a new official emphasis and upgrading of the role of the universities. For decades, the Soviet Academy of Sciences has been routinely hailed as the leader and coordinator of the country's research effort, whether in pure or applied science. While not detracting from this role, Yagodin, however, in an interview on Soviet television, stressed that the universities have a particular contribution to make. In the higher education sector, he said, there is a unique opportunity for interdisciplinary contacts between pure and applied sciences, economics and "the organization of production" — what in the West would be called business studies.

Vera Rich

International space

Space year for Columbus

Washington

MOMENTUM is building up for an International Space Year in 1992. With a strong congressional champion, US participation in the scheme seems assured. What remains uncertain is whether it will be a significant scientific programme, or merely window dressing for the International Geosphere/Biosphere Program, launched last month at the meeting of the International Council of Scientific Unions (ICSU) at Berne and also due to begin in the 1990s.

The idea for an International Space Year, or ISY, came last year from Senator Spark Matsunaga (Democrat, Hawaii), who suggested that it would be a fitting way of marking the 500th anniversary of Columbus's trip to North America, and might prevent the United States from "backing into the Space Age with our eyes

heading for a collision over funds, but that worry seems to have abated, largely because ISY will probably not need new funds but merely the redirection of existing programmes. IGBP will be a big spender on programmes that would continue for as long as a decade.

The selection of 1992 for ISY could be a good choice for more than historical

reasons. Several space missions planned for the 1990s should be in operation that year, including two US and two Japanese satellites. Several Soviet Cosmos satellites are also planned for operation in 1992.

As international bodies begin to consider the implications of joining ISY, plans for IGBP are moving ahead. A secretariat will be established at the Royal Swedish Academy, and appointments to a special ICSU coordinating committee will be named by the beginning of 1987.

Joseph Palca

Japan fellows

British are reluctant travellers

Tokyo

THERE are signs that the Royal Society of London may be able to make fuller use of a scheme for recruiting British postdoctoral fellows to research posts in Japan. That, at least, seems to be the impression left on Japanese speakers at a conference at Bristol and by a visit to Japan earlier this year by Sir Arnold Burgen, the foreign secretary of the Royal Society.

The scheme is organized in Japan by the Japan Society for the Promotion of Science (JSPS) and provides for up to ten fellowships a year to be held by British scientists. In spite of the generous tax-free allowances (amounting to more than £20,000 or ¥4.54 million a year), the Royal Society has hitherto been unable to fill more than about a half of the places available. The cost of the scheme is borne entirely by the Japanese; JSPS officials are anxious that the vacancies be filled, and told Burgen as much when he was here.

Last month's Bristol seminar was planned as the first of a series meant to interest British academics in working in Japan. The speakers included Professor Minoru Oda, director of the Institute of Space and Astronautical Science, an outgrowth of the University of Tokyo, and Mr Hiroshi Kida, director-general of JSPS. There are already substantial links between Japan and Britain in fields such as astronomy, high-energy physics and biotechnology, some within the framework of formal agreements signed earlier in the decade.

In this spirit, the University of Leicester will be providing the proportional counters for the Japanese X-ray astronomy satellite ASTRO-C, due to be launched next year. And agreement is expected soon on the construction by physicists at Tohoku and Tsukuba Universities of a multi-angle rotor spectrometer for the British spallation neutron source at the Rutherford Appleton Laboratory. The Royal Society has also drawn attention to the fruitfulness of collaborations such as that between the Royal Observatory Edinburgh and the Tokyo Astronomical Institute and the annual joint biotechnology workshops.

Why, then, is it so hard to fill the vacant fellowships? Identical schemes run by France and West Germany are regularly over-subscribed. Fear of the unknown seems to be part of the problem, as is uncertainty about future job prospects. Job-hunting from a distance is difficult, while those visiting postdoctoral fellows who might be tempted by teaching posts at Japanese universities discover that such posts are virtually impossible to obtain. The Royal Society is therefore planning to offer one-year fellowships in Britain to people returning from Japan, and will advertise the scheme more widely.

David Swinbanks

UK fund for China

THE visit of the British monarch to China this week, seen as a sign of goodwill over the agreement on the return of Hong Kong to Chinese sovereignty reached earlier this year, was garnished on 13 October by the Queen's announcement in Beijing of a new £1 million fund to allow Chinese postdoctoral scientists to work in UK institutes.

The new fund has been established largely through the efforts of Lord Rhodes, who has raised about one-third of the money from the British government's Overseas Development Agency and the rest from about 20 of the larger companies operating in the United Kingdom. The Royal Society will run the fund and will offer the first of 30 one-year awards that will be known as Royal Fellowships at the start of next year in collaboration with the Chinese Academy of Sciences and the China Association of Science and Technology.

There is an increasing demand from the Chinese for the opportunity to work in foreign laboratories. About 20,000 research workers and students are currently abroad, mostly in the United States, Japan and West Germany. Only about 1,400 are in the United Kingdom, which has lagged behind in the provision of places, but it is hoped the new fund will mark an upturn and stimulate the provision of further schemes.

Nigel Williams



fixed firmly on the ground". At the request of Congress, the National Aeronautics and Space Administration (NASA) formed an interagency group to look into the proposal and reported in May that the international response to the idea was "overwhelmingly positive". NASA suggested that the Committee for Space Research (COSPAR) might be an appropriate organizing body. In July, COSPAR decided to form a preliminary committee to coordinate preparations for ISY, using the International Geophysical Year as a model.

The National Research Council, the research arm of the National Academy of Sciences, met in August to discuss its role in ISY. Just last month, COSPAR's parent body, ICSU, decided to form an *ad hoc* committee to investigate the plan. ICSU will next October make a final decision whether formally to join ISY. The International Astronautical Federation is also considering its role in ISY.

Supporters of ISY say it will be a significant worldwide educational effort, encouraging a better understanding of space science. For a time, it appeared that ISY and the International Geosphere/Biosphere Program (IGBP) might be

Direct broadcasting

Unravelling an encryption code

A TECHNICAL standard for encoding European television could be on the way. A study of the technical and economic aspects of subscription television is to be carried out in the next six months by the London-based consultancy group Communications Studies and Planning International (CSPI), through a commission from the British Home Office.

The study is an attempt to comply with one of the findings of a report on the financing of the British Broadcasting Corporation (BBC) published in July and compiled by a committee led by Professor Alan Peacock, research professor in public finance at the Esmée Fairbairn Research Centre, Heriot-Watt University in Edinburgh. It recommended that subscription television might be an option for the future financing of the BBC. The results of the CSPI research, however, will have wider implications.

Direct broadcasting by satellite (DBS) television is to be launched in mainland Europe next year, supposedly by West Germany in the spring, followed by France in the autumn. By the end of this year, the British are to select an operator for a DBS television service across the United Kingdom. A potential audience of 120 million European homes could receive television programmes via a satellite dish antenna about 90 cm in diameter. In most cases, the pictures will be encrypted and the service financed by subscription.

How the British DBS television service will operate is still unclear. The Independent Broadcasting Authority (IBA), the controlling authority over British commercial broadcasting, has been charged with the task of selecting the British DBS operator from five applicants. The IBA decision is expected by the end of the year and would mean the launch of a British satellite by about 1990.

The British are thus at least two years behind West Germany and France in their commitment to DBS. The vacillations of the BBC, which was originally awarded the DBS franchise about four years ago, caused much of the delay. The corporation proved reluctant at that time to commit itself to an expenditure estimated to be more than £350 million over a seven-year life of a DBS satellite.

The CPSI research will cull a great deal from the US experience where a plethora of cable television networks has meant the extensive use and development of encryption and subscription techniques. The result could be a single encryption/subscription technical standard in the United Kingdom for terrestrial and satellite broadcasting.

Bill Johnstone

Nobel prizes

Growth factors bring rewards

WITHIN weeks of winning this year's Lasker award in basic science (see *Nature* 323, 289; 1986), Stanley Cohen of Vanderbilt University School of Medicine in Nashville and Rita Levi-Montalcini of the Institute of Cell Biology in Rome have shared the 1986 Nobel Prize in Physiology or Medicine. Again, the citation is for their work on growth factors. Levi-Montalcini has spent much of her life working on nerve growth factor (NGF), which she dis-



Levi-Montalcini (left) and Cohen. Lasker awards last month, Nobels this.

covered in the early 1950s; Cohen worked with her at Washington University, St Louis, on the purification of NGF between 1956 and 1958 but on moving to Vanderbilt turned his attention to epidermal growth factor (EGF), upon which he has continued to focus.

Remarkably, Levi-Montalcini was a medical student in Turin, Italy, with two other subsequent Nobel laureates, Salvador Luria (Physiology or Medicine 1969) and Renato Dulbecco (Physiology or Medicine 1975). She graduated in 1936 but left Italy for Brussels in 1939 after Jews were barred from academic or professional careers. Returning to Italy shortly before Belgium was invaded by Germany, she built a laboratory in her bedroom and set about studying the effect of peripheral tissues on the development of nerve centres. In this she was joined by Professor Giuseppe Levi, previously her teacher.

Levi-Montalcini has attributed her devotion to a small neuroembryological problem "when all the values I cherished were being crushed" to "the well-known refusal of human beings to accept reality at its face value". Forced to leave Turin because of heavy bombing in July 1942, she rebuilt her laboratory in a small country house, under far worse conditions. Electric power was cut off every few days and the eggs she used for her experiments were in short supply. A year later she had to flee to Florence where she assumed a false identity and a precarious existence until the end of the war.

Fortunately, a paper by Levi-Montalcini and Levi, rejected by Italian

journals on account of the non-Aryan names of its authors, had been published by the Belgian *Archives de Biologie*, and read by Viktor Hamburger, who in 1946 invited Levi-Montalcini to join him in St Louis.

By 1951 they had discovered that certain mouse tumours release a substance that induces nerve fibres to grow. The bioassay for the substance was developed during a visit of Levi-Montalcini to the Biophysics Institute of the Medical School of Rio de Janeiro in 1952 and the substance assumed its present name in 1954 — having been termed nerve growth stimulating factor until then.

Shortly thereafter, Stanley Cohen joined Levi-Montalcini in an attempt to purify NGF and in the course of so doing discovered that snake venom and mouse salivary glands were rich sources of the factor. Cohen identified NGF as a protein and established its molecular weight. He also made antisera to purified NGF and showed that they would destroy the sympathetic ganglia of newborn mice.

Levi-Montalcini returned to Italy in 1961 and has since worked in Rome. Much of her work has been on the functions and mode of action of NGF, which was finally sequenced in 1971.

On moving to Vanderbilt, Cohen began his studies of another factor in salivary glands, which he isolated in 1962 by its ability to accelerate eyelid opening in the newborn mouse, still the standard bioassay. Cohen had determined the structure of EGF by 1972 and thereafter increasingly turned his attention to the factor's receptor.

He established that after EGF has bound to its cell surface receptor, the complex is internalized into the cell (one of the first examples of receptor-mediated endocytosis, a central component of the work for which Michael S. Brown and Joseph L. Goldstein shared the Nobel Prize for Physiology or Medicine last year), and that the complex has tyrosine protein kinase activity, which provides a link to oncogenes.

Cohen is much admired for his modesty and the fact that he still gets his hands dirty in the laboratory. He is, says an admirer, an experimentalist *par excellence*, who has always been first and right in his experiments. Of his time spent working with her on NGF, Rita Levi-Montalcini has written "Stan used to spend the entire day and most of his evenings ... meditating with eyes half-closed smoking his pipe, and playing his flute (his main talents were, however, not in this direction)". The Columbus Day and Yom Kippur holiday this week prevented a progress report on his flute-playing.

Peter Newmark

Asylum

US cancer researcher in Moscow

Washington

To the astonishment of his former co-workers, Arnold Lockshin appeared on Soviet television last week saying that he had defected from the West. Lockshin had been on the staff of the Stehlin Foundation for Cancer Research in Houston, Texas, until he was fired on 19 August. Stehlin officials say Lockshin was dismissed because of poor job performance.

Lockshin was awarded a doctorate in biochemistry from the University of Wisconsin in 1966. His dissertation research was on photosynthesis. After moving to the Boston area he collaborated with Lawrence Bogorad at Harvard on work on chloroplasts. He became involved in cancer research while working at the University of Southern California. In the early 1980s, he joined the Stehlin Foundation, working chiefly on the effects of various cancer therapies on human tumours expressed in nude mice. He is the senior author of 20 scientific articles.

Most recently, Lockshin has been concerned to set up animal models for the study of the mechanism of immunosuppression by cyclosporin A, and the formation of tumours as a consequence. Some of Lockshin's recent work in this field was submitted for publication in *Nature* earlier in the summer. Colleagues in Texas say that the letter saying that the article concerned would be more suitably placed in a specialist journal arrived in Houston the day that Lockshin was fired.

The laboratory director of the Stehlin Foundation, Beppino Giovannella, says that Lockshin gave no indication of his intention to defect, but he acknowledges that Lockshin's political views were "not exactly right wing". Robert Burris, Lockshin's graduate adviser at Wisconsin, says Lockshin visited Cuba while a graduate student. First reports from Moscow had Lockshin's name spelled Lokshin, causing

confusion about his identity.

Lockshin's own account of his defection is somewhat different. He said on television that since the time of the Vietnam war he and his wife had been active in "the anti-war movement, the movement for peace and for changes in the social policies of the United States". As a result, he claimed, he and his family suffered "endless threats by telephone and letter", and also "pogroms" against his laboratory and apartment. His dismissal, he said, was on political grounds. He hopes to resume his research work, in the Soviet Union, he said, in due course.

In back-up interviews on the foreign services of Soviet radio, Lockshin amplified his story. His latest political action, he said, was a demonstration in Houston against President Reagan's Strategic Defense Initiative. The threats against his wife included, he said, an "ex-green-beret waving a gun in her face", while he himself had been threatened with "physical annihilation". The US "political police", he alleged, had intercepted his mail, tapped his telephone calls and generally orchestrated the campaign of harassment and denigration.

The Soviet media, not surprisingly, made the most of the incident, particularly as it came just before the Reykjavik summit. Some commentators, however, clearly over-reacted — Aleksandr Barabichik of Moscow Radio's "World Service" construed the fact that the Americans had failed to identify Lockshin (because of the misspelling of his name) as "a kind of sociological warfare". And none of the commentators mentioned the fact that, in spite of the alleged harassment, Lockshin was able to contact the Soviet Embassy in Washington, arrange his political asylum and book flights to Moscow for himself, his wife and three children without let or hindrance.

Joseph Palca & Vera Rich

"Extinct" lemurs found in Madagascar

Washington

A GROUP of about 35 greater bamboo lemurs (*Haplolemur simus*), feared to have been extinct since the early 1970s, has been found in a remote rainforest in southeastern Madagascar.

The animals were discovered by Patricia Wright of Duke University during a recent expedition. The fate of the species is far from assured, however, because unless there are other surviving groups, the genetic variation present may be too low for recovery.

Furthermore, the animals, which feed on bamboo, are not found in any of the 11 wildlife reserves on the island. Efforts are



now under way to persuade the Malagasy government to institute a special reserve for the lemurs. □

British research

Better measures of decline

THE Advisory Board for the Research Councils (ABRC), interposed between the grant-making research councils and the British government, should now be better placed to chart the decline of British basic science. That seems to be the chief message of two reports commissioned by the advisory board and published this week.

An account of the findings of the first report prepared by its authors, Ben Martin of the Science Policy Research Unit at the University of Sussex and John Irvine of the Technical Change Centre in London, appears in this issue on page 591.

The second document, the result of research by the Science and Engineering Policy Studies Unit, jointly run by the Royal Society and the Fellowship of Engineering, consists of a detailed comparison of the productivity of various countries in some specific fields of research. An extended summary of that document by its authors will appear in next week's *Nature*.

The publication of the two reports is something of a landmark for ABRC, whose activities have been traditionally confined to advising the British government on the proper division of the research budget among its potential recipients.

Although this advice has been published on some occasions, ABRC has not previously commissioned and published studies of particular aspects of British administration although its predecessor, the Council for Science Policy (disbanded after the Rothschild report of 1971) regularly undertook studies of the administration of science. ABRC now confidently labels its document "No.1", and promises that there will be more to follow.

To some extent, the studies now published overlap. Each, for example, is concerned with the scale of government spending on civil science; each concludes that, by different yardsticks, Britain lags behind its industrial competitors in Western Europe as well as Japan and the United States.

The Royal Society document is chiefly concerned to test the applicability of "bibliometric indices", citation indices and the like, as measures of the output of research. Interestingly, these techniques have been zealously applied to a number of research fields in which British scientists are active by the authors of the first report, Martin and Irvine.

Even though cautiously, the authors of the Royal Society report conclude that bibliometric analysis is a useful measure of a research community's standing relative to others. □

European research

Joint centres in eye of storm

A PROMISED reform of the European Commission's big spenders in research, the Joint Research Centres at Ispra in Italy, Karlsruhe in Germany, Petten in the Netherlands and Geel in Belgium, is raising hackles in European standards institutions, because it may threaten their work in standards and measurements. But according to the new director of the research centres, Jean-Pierre Contzen, "we don't intend to tread on any flowerbeds".

The Commission is nevertheless treading a narrow path between governments, which wish to see better value for money from the 200 million ECU (£150 million) they spend each year on the research centres (largely a hangover from the boom years of nuclear energy research), and the national laboratories nervous of a challenge to their roles. Advisers have identified the setting of Europe-wide standards in new composite aeronautical materials, for example, as a logical job for the research centres. But such work could mean competition for national institutions wishing to move away from their classical task of defining the units of physical measurement such as the metre and the second and towards a more secure role involving links with industry.

Contzen, however, is assuring national laboratories that he aims to find a place for the research centres that will complement national work. Work at the centres will be highly applied. They could be involved in the strong political effort to reduce trade barriers in Europe by establishing themselves as an objective, impartial European centre of reference in squabbles over incompatible technical regulations (over consumer electronics connections and safety requirements, for example). The centres could also take on the research necessary for setting common standards in new industrial processes and technologies, particularly in matters relating to European environmental regulations, Contzen says.

The expertise of the joint research centre staff is at present mostly in nuclear matters. According to Contzen, Karlsruhe excels in the study of actinides, Geel in nuclear measurement and Ispra in nuclear safety and the assessment of the risks of various forms of nuclear waste disposal. But Petten has done research in high temperature materials since 1976, and is now working on ceramic magnets. Ispra also has a background in structural mechanics; and was chosen by the US National Aeronautics and Space Administration and the European Space Agency as a remote sensing centre. Here Ispra's interests are in developing applications in agriculture, oceanography and in tackling the problems of poorer countries. Ispra also runs

the European solar testing installation for investigating new solar energy technologies, makes engineering and safety studies for nuclear fusion and gives support to the planning for the demonstration-scale Next European Torus (NET) that should follow the Joint European Torus (JET). But Ispra's environmental work on lead in blood and toxic chemicals has been "more controversial", Contzen says.

It is this team that Contzen now hopes to turn more towards the interests of member states. He is awaiting the advice

of a panel of European industrialists on improving the research centres' value to industry. This advice, however, is not due before the December meeting of the European Council of Ministers which should fix the outlines of Commission research spending to 1991. The Commission has allowed for a 5 per cent fall in research centres' spending next year, and with the research centres' background in nuclear safety has probably been protected from pressure for greater reductions by the event at Chernobyl. But uncertainty about the research centres' exact programme will remain, and with it the concerns of the national laboratories that still feel themselves threatened.

Robert Walgate

Japanese astronomy

Projects wait on organization

Tokyo

JAPAN'S Ministry of Education, Science and Culture has requested a preparatory budget from the Ministry of Finance to reorganize the Tokyo Astronomical Observatory of the University of Tokyo as an independent national institute. The plan would involve a merger with at least one other astronomical institute, the International Observatory in Mizusawa in northern Honshu, and is closely linked to several new big projects proposed by Japan's astronomers.

Tokyo Astronomical Observatory was founded in 1878 and moved to its present site in Mitaka in the suburbs of Tokyo in 1924. It now operates six other observatories around the country, including a 188-cm reflecting telescope at Okayama, a 105-cm Schmidt telescope at Kiso and the 45-m high-frequency radio telescope at Nobeyama. The International Latitude Observatory at Mizusawa goes back to 1899, and is concerned primarily with geodetic observations of latitude, longitude and Earth rotation.

Moves by the government to reorganize Tokyo Astronomical Observatory first emerged publicly with the observatory's proposal to build a giant 7.5-m optical and infrared telescope at Mauna Kea in Hawaii (*Nature* 311, 5; 1984 and p.515, this issue). The project was simply too big and too international for a mere branch of Tokyo University to run.

Although the national committee of astronomy of the Science Council of Japan last year recommended the building of the Hawaii telescope and a plot of land has already been reserved from the University of Hawaii, the Ministry of Education, Science and Culture has yet to be persuaded to submit the telescope proposal to the Ministry of Finance. The option on the land in Hawaii runs out in 1988, by which time Tokyo Astronomical Observatory should be a national institute, if all goes according to plan.

A couple of other big projects lie behind the reorganization. Geodetic astronomers, including those at Mizusawa, wish to study the rotational motion of the Earth by building a very long baseline interferometry network including satellites. And it is hoped that the same net can be used by radioastronomers for their own purposes.

In turn, solar radioastronomers at Nobeyama and the Research Institute of Atmospherics at Nagoya University have proposed building a giant radio-heliograph with 200 dishes. Although the Research Institute of Atmospherics is not yet officially involved in the reorganization plan (the institute already has a restructuring programme), its solar radio group is likely to join the new central institute to be organized around Tokyo Astronomical Observatory.

Two major issues have yet to be resolved. First, Tokyo University, which has its own astronomy department separate from the observatory, is not prepared to part completely with the observatory. Two or three sections and one telescope (possibly the 188-cm reflector at Okayama) will remain within the university. But which sections will remain has yet to be decided. A similar situation arose during the formation of the Institute of Space and Astronomical Sciences (ISAS). When it separated from Tokyo University, some personnel remained behind.

Second, the new independent institute will have to find a new source of graduate students. The Ministry of Education, Science and Culture is planning to establish a new graduate university to supply the needs of national research institutes, such as the Institute of High Energy Physics at Tsukuba. But Tokyo Astronomical Observatory would prefer to continue to draw students from Tokyo University in the same way as is done at ISAS, and this would fit in with the plan to retain some sections of the observatory within the university.

David Swinbanks

US astronomy

Big ideas, short pockets abound

Washington

LARGE optical telescopes that are also usable in the infrared are all the rage these days, but paying for them is another matter. With the exception of the Keck Telescope, a joint project of the California Institute of Technology and the University of California, large telescopes are being planned, not built. But now hope flickers for an ambitious project called the Binocular Telescope.

What has happened is that the first sponsors of this project, the University of Arizona and Ohio State University, have found a third partner, the University of Chicago, to join them in building a twin 8-m telescope. And now there is also a chance that a fourth partner, a consortium of Italian institutions, will join in. But even with four partners, the binocular telescope will exist only on paper until the estimated \$60 million needed to build it can be raised.

The instrument will consist of two 8-m mirrors side by side on the same mount. The twin mirror design will give the light-gathering capacity of a 11.3-m telescope and the resolution of a 22-m telescope. Peter Strittmatter, director of the Steward Observatory at the University of Arizona, says the new telescope will have an angular resolution of 5 milli-arc seconds in the visible spectrum.

A site selected for the telescope on Mount Graham in southeast Arizona was given preliminary approval by the US Forest Service last week. The mirrors will be made using a new spin-casting technique being developed by Roger Angel at the University of Arizona. A 3.5-m mirror blank will be cast early next year to test this new technique. The plan is for the telescope to be operating by 1992, but that may be optimistic.

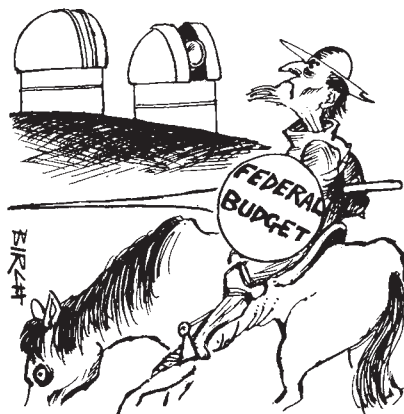
The University of Chicago's decision to become a partner was not unexpected. Astronomy department chairman Roger Hildebrand says Chicago has had an historical interest in large telescope projects. In the 1890s, Chicago built the Yerkes Observatory and, in the 1930s, the McDonald Observatory in Texas. The astronomy department is now especially anxious to work with optical interferometry techniques, for which the binocular telescope offers many opportunities. Chicago has also recently joined a consortium to build a remotely operated 3.5-m telescope in New Mexico, a project that will also use a spin-cast mirror from the University of Arizona (see *Nature* 322, 493; 1986).

Several US telescope projects are now being canvassed. The Carnegie Institution of Washington has begun discussions with the University of Arizona and Johns Hopkins University on a Southern Hemis-

phere telescope. Located at Las Campanas in Chile, that would have a single 8-m mirror. The University of Texas also has plans for a single-mirror 8-m telescope to be built in Texas.

Interest in new large telescopes is not confined to the United States. Japan is planning an 8-m telescope, similar to that being considered by Texas, which would be located at Mauna Kea in Hawaii, while the European Southern Observatory (ESO) has plans for a telescope consisting of four 8-m mirrors, to be located in Chile.

The ESO project is very similar to that for the National New Technology Telescope (NNTT) being considered by the US federal agencies. That would also con-



sist of four 8-m mirrors to be cast and polished using Angel's techniques. The selection of a site for NNTT is still under way, but it will definitely be in the Northern Hemisphere. As a national project, NNTT does not have the luxury of seeking private funds, and it is thought to be "far down in the queue" for new starts by federal agencies. Ironically, many of the large telescopes under consideration are likely to use mirrors of the type being developed at Arizona for NNTT.

Although paper and pencil (and computer) are still the primary tools for building most new large telescopes, project managers of the Keck Telescope have switched to bulldozers and glass furnaces. Construction of the telescope housing and mount has begun at Mauna Kea.

Forty-two mirrors are being made for the Keck Telescope, which will be a composite parabolic surface based on 36 separate mirror surfaces. There will be one spare for each of the 6 partial parabolas used in the segmented mirror design. Earlier this year, project directors successfully tested the intricate pointing and stabilizing system the Keck Telescope will use to control its multiple mirrors. First light is expected in 1990, although the project still needs to raise an additional \$17 million to meet the final expected production cost of \$87 million.

Joseph Palca

French research

Optimists and pessimists

M. ALAIN Devaquet, French minister for research and higher education, "appeared to be a happy man" when presenting his research policy in Paris last week, in the wake of the recent 1987 budget announcement, according to a reporter from *Le Monde*. Less happy, however, was the panel of "wise men", the Conseil Supérieur de la Recherche et de la Technologie, which saw fit the same day to publish its warnings on the future of French science, the panel's advice to the minister first delivered in secret on 3 July.

That the Conseil Supérieur, consisting of scientists, trade unionists and industrialists, should believe its warnings of three months ago to be still appropriate suggests that despite his smiles, Devaquet has not yet set the fears of the French scientific and technological community at rest. There have been improvements since July, but they came mostly in the form of a dramatically increased defence budget. The consequent extra FF5,000 million (£500 million) for research and development will be spent mainly in industry, where it will do something to offset cuts in industrial research subsidies elsewhere in the budget; but the spending will be only in a few big industries, the Conseil complains. Moreover, the ministry of defence has still not announced the areas of military research on which its bonanza will be spent, so it is still difficult to assess the overall impact. Research in small companies and related long-term market share and exports, are still threatened, the Conseil says.

In its earlier advice now published, the Conseil complained that:

- The loss of 373 research posts was the first reduction in French research personnel "since the Second World War".
- Basic research in 1987 would be supported only with the same real budgets as in 1985, despite increases in research costs. The re-equipment of French laboratories, previously planned for 1986, remains far from complete.
- The balance between the French "big programmes" (space, nuclear power and aviation), industrial research and strategic and basic research has gone too far in favour of the big programmes.
- Research support removed from small companies would prove "irreplaceable".

For Devaquet there must be one consolation, however. The present government has placed almost all its ministries under tight financial constraint, and yet the research minister has achieved an increase in his budget. Although this merely restores spending to 1985 levels, the political achievement is real. Robert Walgate

Grading of universities

SIR—While there is much evidence that many university departments were wrongly graded by the University Grants Committee (UGC) earlier this year (see *Nature* 322, 299; 1986), little attention has been

middle third.

Statistical analysis shows that, while there is a significant correlation between the two methods of assessment ($p < 0.01$), the measures of research activity I have

Percentage distribution of ranks (pooled data).

	*	A+	A	A-
Biological science	8	26	35	31
Physical science	9	28	26	38
Computing science	12	18	24	47
Mathematics	25	17	34	25
Social science	26	11	37	27
Arts	20	26	45	8

given to UGC's use of the term "average", the meaning of which appears to have varied from one committee to another.

The table shows the number of different departments placed in the four UGC categories: the starred departments are those whose research was deemed to be of international standard. It will be seen that the distribution of departments among the four categories is odd. For example, in the Arts, the chance of a department being below average was only 1 in 12, but in computing it was 1 in 2.

The data show that it is dangerous to use the UGC assessments for a comparison of departments in different fields (or "cost centres"), yet many universities are already doing just this, while UGC is planning to distribute its special equipment grant on the same basis. Earlier criticisms of the assessment procedures have been met with statements to the effect that UGC consulted everybody it could think of, so what else could it have done? In practice, it seems that the information thus gathered may have been less important than the manner in which it was used, with some committees "coming to a conclusion" by hunch. On such a potentially important matter, there should surely have been a formal model for assessment, with the various factors stated explicitly and weighed properly.

I have constructed such a model for physiology departments at British universities. For the assessment of research activity, I used four measures: total research grant income, total grant income from the research councils, total PhD students and the numbers of PhD students allocated to departments under the research councils' quota system. Departments' scores under the four headings were divided by the numbers of permanent staff and then ranked in numerical order. Departments in the middle third of each distribution, counted as "average", were allocated 2 points and those below and above 1 and 3 points respectively. The sum of the scores under the four headings in principle ranging between 4 and 12, were again divided into three groups, with that labelled "average" consisting of the

used account only for about 40 per cent of the UGC grading, 16 per cent of which appears to have been determined by the size of departments (the larger the size, the higher the grade). It would be interesting if other research areas in British universities could "test the assessors" by means of similar data.

J.F. LAMB

*Department of Physiology
and Pharmacology,
St Andrews University,
Fife, Scotland KY16 9TS, UK*

Defining darwinism

SIR—My colleague M. Sinclair, in his letter¹ on the Imanishi-Halstead issue², correctly insists that without intraspecific competition, any evolutionary theory must be considered to be "non-darwinian"; the baby would be thrown out with the bathwater. I wish to take up another aspect of Halstead's argument, that which maintains that Imanishi's "attack" is "merely a caricature of Darwin".

The categorization of anything that disagrees with "modern darwinism" as a misreading, a misunderstanding, a failure to appreciate the full subtleties of the position or, in this case, as bordering on *lèse majesté*, I find distasteful; I regret to have to say that this is not an isolated example either on Halstead's part³ or, more generally, on the part of the average member of the neo-darwinian priesthood⁴.

The very title of Halstead's essay, "Anti-darwinian theory in Japan" is provocative, with its implication of a purely negative attitude on the part of Imanishi. This leads me to my suggestion, framed as a question: why should the study of evolution be called "darwinism"? Was Einstein "anti-newtonian"? Has the development of modern physics, beyond anything that Newton could possibly have conceived in his day, detracted in the slightest from his stature?

"Darwinism" and "neo-darwinism" are touted by the faithful as the great unifying concepts in modern biology — which is ludicrously untrue. They are concepts as unifying as Christianity, Mohammedan-

ism or the general belief in a deity. They serve less to perpetuate Darwin's honoured place in science than to justify the claims of acolytes to orthodoxy. They are not as useful as guides to research workers as is claimed and can be excuses for sloppy thinking and the stifling of promising alternatives.

When a graduate student, I ate many meals at Christ's College, Cambridge, under Darwin's eye. Possibly, I now risk excoriation for suggesting that the memory of perhaps the most distinguished alumnus of the college should be perpetuated differently. In justification I invoke the memory of another pre-eminent member of the same college and suggest that *Paradise Lost* could be followed by *Paradise Regained*; Darwin's memory should be neither a crutch nor a club.

T.D. ILES

*Fisheries Research Branch,
Scotia-Fundy Region,
Biological Station, St Andrews,
New Brunswick EOG 2X0, Canada*

1. Sinclair, M. *Nature* 320, 580 (1986).
2. Halstead, B. *Nature* 317, 587-589 (1985).
3. Halstead, B. *New Sci.* 100, 940-941 (1983).
4. Ridley, M. *Nature* 318, 124-125 (1985).

Dioxin risk

SIR—We object to your publication of the letter by Drs Alastair Hay and Ellen Silbergeld ("Assessing the risk of dioxin exposure", *Nature* 315, 102; 1985) because they do not disclose their "hidden affiliation".

Both Hay and Silbergeld have been paid consultants to lawyers representing plaintiffs in a lawsuit against Monsanto Company. Although both Drs Hay and Silbergeld testified in court under oath, they did not see fit to put their analysis as described in *Nature* into the record.

ALLAN M. FORD
W.J. MCCARVILLE

*Monsanto Company,
800 N. Lindbergh Boulevard,
St Louis, Missouri 63167, USA*

Hay and Silbergeld reply: It is true that we appeared as expert witnesses in a lawsuit against Monsanto. We criticized Monsanto epidemiology studies while under oath. We regret that Ford and McCarville do not comment on the substance of our critique.

ALASTAIR HAY
*Department of Chemical Pathology,
University of Leeds,
Leeds LS2 9NL, UK*

ELLEN SILBERGELD
*Environmental Defense Fund,
1525 Eighteenth Street, NW,
Washington, DC 20036, USA*

● Contributors who fail to disclose affiliations are in danger of being thought to mislead their readers — Editor, *Nature*.

The reality of the quantum jump

Three-quarters of a century after the concept of quantum jump was introduced by Bohr, three independent measurements have emerged on each other's heels.

THE phrase "quantum jump" has long since escaped from physics; as a jump or, more dramatically, a "leap", the concept is applied in a variety of unconnected fields, from political upheavals to the emergence of new techniques in, say, modern dance. Quite properly, quantum jump has become a descriptor of discontinuity. So it may be of some general interest that quantum jumps, in the sense in which Bohr first used the term more than half a century ago, have just been observed for the first time.

Why should it have taken so long? The basic idea is that an atom capable of absorbing a light quantum of some determined frequency will, when it does so, jump into another state "instantaneously". The adverb of time is not to be taken literally, but guessed at from Heisenberg's uncertainty relations, but implies that the time taken to jump is a small fraction of the periodic time corresponding to the frequency of the radiation.

So the experiment to demonstrate that quantum jumps are real is simply designed: take a single atom, arrange that it absorbs a photon whose energy corresponds to the difference between one electronic state and another and then demonstrate that the atom does jump from one state to another as the photon is absorbed.

But only in the past few years have people been able seriously to think of "taking a single atom"; for several reasons, but chiefly because of the opportunities for accurate spectroscopy provided by well-tuned laser beams, there has evolved a family of electromagnetic traps for single atoms. The other ingredients in this recipe for an experiment are not so simple either; demonstrating that there has been a quantum jump will simple-mindedly entail a measurement of some kind which, if badly planned, will disturb the state it is intended to demonstrate. So the three successful experiments now independently reported are technical as well as intellectual achievements.

The essence of the two experiments is the technique for showing that there have indeed been jumps, which appears most immediately to derive from a suggestion last year by Richard J. Cook and H. J. Kimble (*Phys. Rev. Lett.* **54**, 1023; 1985), who in turn credit the University of Washington physicist Hans Dehmelt a decade earlier.

The trick is to choose an atom with an electronic state from which there are two well-understood transitions to two diffe-

rent excited states, one of them readily accomplished (strong) and the other less so (weak). Irradiating a single atom with a laser whose frequency corresponds to the energy of the strongly-coupled transition has an easily predictable effect. The atom is repeatedly excited and, spontaneously, de-excited; the fluorescent radiation is emitted in all directions, and is not coherent with the laser beam, so that the atom is a powerful scatterer of radiation. Arranging that a single atom will scatter as many as 10^6 photons a second appears not to be particularly difficult.

If the same atom is also irradiated with a laser corresponding to the energy difference of the weakly coupled transition, there will also be occasions when the atom is excited into the weakly coupled, not the strongly coupled state. And those occasions will be recognizable because the atom will no longer be fluorescent. Cook and Kimble were chiefly concerned to analyse the statistics of these occurrences, showing that they would form a pattern characteristic of Bohr jumping and also serve as a means of measuring the spontaneous transition probabilities between weakly coupled electronic states.

The three experiments now reported have the virtue of relying on two different atoms with very different transition schemes. A group from the University of Hamburg (Sauter, Th., Neuhauser, W., Blatt, R. & Toschek, P.E., *Phys. Rev. Lett.* **57**, 1696; 1986) have worked with singly ionized barium atoms, as have Warren Nagourney, Jon Sandberg and Hans Dehmelt (*Phys. Rev. Lett.* **56**, 2797; 1986).

In each case, the competing weak and strong transitions are decays from an excited state to alternative de-excited states. But J. C. Bergquist, R. G. Hulet, Wayne M. Itano and D. J. Wineland (*Phys. Rev. Lett.* **57**, 1699; 1986) from the US National Bureau of Standards at Boulder, Colorado, have worked with a set of three energy-levels of singly ionized mercury atoms in which the competing weak and strong transitions are excitations, replicating the case described by Cook and Kimble. The spectroscopic details are not strictly relevant to the essence of the result; all that matters is that the weak transition should, from time to time, temporarily sterilize the strong transition.

Inevitably, expectations are confirmed. The traces of the detectors measuring the fluorescence radiation are indeed interrupted from time to time, for periods

whose duration is a measure of the strength of the weak transition. The three sets of experiments confirm that the time distribution of these intervals is exponential, as it should be. It is especially intriguing that, when two mercury ions are bundled together in the same electromagnetic trap, the fluorescence may occasionally be quenched not by the full amount but by exactly half, corresponding to a weak transition in only one of the ions.

At this late stage, with the notion of the quantum jump so firmly in charge of people's minds, the confirmation that quantum jumps actually occur is bound to seem a somewhat academic matter. But that entirely neglects the value of systems such as those developed for these measurements not merely in measuring the lifetimes of weakly interacting states (in which astrophysicists have a vivid interest) but of unravelling the various processes that contribute to the de-excitation of these states; in some circumstances, interatomic collisions matter most, in others, stimulation by radiation of the right frequency may determine the rate.

The way in which these experiments have evolved is, at the same time, a telling proof of the stimulus of new instrumental developments. Electromagnetic traps for single ions, or small numbers of them, were an obviously desirable technique five years ago. The spectroscopists were the chief customers. But the development of these instruments has lent credibility to the manipulation of single atoms, and the consequence of that has been a string of intriguing experiments in basic physics, the demonstration of quantum jumps among them. Perhaps the most intriguing so far is the demonstration, a few years back, that the transition probabilities between the energy levels of an atom placed in a conducting cavity are affected by the shape and size of the cavity, which is a direct but qualitative proof that quantum field theory is everything that people claim for it.

That, in all these cases, the direct demonstration of a principle follows by several decades the assumption that it is correct should not be a surprise. Most discoveries of principle, or at least the important discoveries, are inferences whose importance derives from knowing that demonstration is impractical. When the demonstrations can be carried out, the principles have been accepted. Galileo's alleged experiment at Pisa illustrates the point.

John Maddox

Perception

Questions of sex and colour

from John D. Mollon

"WHEN I look at that purple clematis, do I have the same colour sensation as you?" When this familiar question is put to him, the professional colour scientist will deftly refer the questioner to Wittgenstein. But there is a more tractable question that he will discuss for hours: if we take any two healthy people (excluding those who are explicitly colour blind) and if we find two different spectral mixtures that look the same to observer A, then will the mixtures always match for B?

On page 623 of this issue¹, Jay Neitz and Gerald Jacobs report that there are two distinct types of men within the population that we call 'colour normal', and that there are three types of women. The authors postulate a polymorphism at the locus on the X chromosome that specifies the long-wave (red-sensitive) photopigment of the retina. Because men are XY, they can inherit only one or other of the two alleles envisaged by Neitz and Jacobs. There should be homozygous women who resemble each of the male hemizygous types; but a third class of women will be heterozygous at the locus and should exhibit an intermediate form of colour vision.

The perceptual test used by Neitz and Jacobs is a variant of one described in *Nature* in 1881 by Lord Rayleigh²: the observer is shown a patch of monochromatic orange light (the sodium line at 589 nm in Rayleigh's experiments) and is invited to match it by adjusting the ratio of red to green illuminating a second patch. Provided he is allowed to adjust, if necessary, the radiance of the orange field, the colour-normal observer is able to make the two fields look identical. He can do this because the short-wave cones of the retina are very insensitive in this part of spectrum and the observer needs only to adjust the red/green ratio to give a ratio of quantum catches in his long- and middle-wave cone receptors that is the same as the ratio produced by the orange light³.

Neitz and Jacobs' test differs in at least three ways from the classical 'Rayleigh match': (1) the stimulus field is large, with an outer diameter of 11 degrees and an occluded foveal region; (2) the fields to be matched are not adjacent but are substituted for each other in time; and (3) the monochromatic field has a wavelength of 600 nm and the red primary is at the very long wavelength of 690 nm.

In the history of colour science, bimodalities among colour-normal observers have often been suggested, and often denied. There have been recurrent reports of a bimodality in the wavelength

judged to be 'unique' green, the green that is neither bluish nor yellowish^{4,5}; but this bimodality has proved difficult to replicate^{6,7} and the physiological basis for the judgement is uncertain.

More relevant to the new results is the



The 'anomaloscope' used by Lord Rayleigh in 1881 to study individual differences in colour matches. This photograph was taken in 1973 while Rayleigh's instrument was being examined by two distinguished scientists, M. Alpern (left) and the late W.A.H. Rushton.

work of Georg Waaler^{8,9}, a retired Norwegian forensic scientist who also measured Rayleigh matches. Using the Nagel anomaloscope (Model II), a clinical instrument, Waaler obtained a series of matches for different wavelengths of the monochromatic field between 574 and 603 nm. He claimed that his male observers fell into two clear groups, G₁ and G₂, according to whether they used more or less red in their matches; occasional matches might be out of line, but the set of equations for an individual allowed a clear diagnosis. Postulating the same mode of inheritance as do Jacobs and Neitz, Waaler showed that G₁ sons never have G₂ mothers and G₂ sons never have G₁ mothers. The new results of Jacobs and Neitz should cause us to look afresh at Waaler's long-neglected claims.

Neitz and Jacobs reported their remarkable results at a meeting of the Optical Society of America last year. At about the same time, visual scientists began to learn of the work of Jeremy Nathans and his collaborators on the molecular genetics of colour vision (refs 10, 11; see my previous *News and Views* article¹²). Nathans' group showed that the X chromosome often carries more than one copy of the gene for

the middle-wave pigment, but there is never more than one copy of the gene for the long-wave pigment. Neitz and Jacobs, on the other hand, postulated variation in the long-wave gene. Many saw here a clear inconsistency.

In fact, there is no contradiction in the results so far published. Nathans and his colleagues sequenced only the two middle-wave genes of Nathans' own X chromosome. The two sequences differ only in a 'silent' substitution of one nucleotide, a substitution that would leave unchanged the sequence of amino acids in the photopigment molecule specified by the gene. In principle at least, it is possible that the middle-wave gene varies in number and the long-wave gene varies in sequence. The conclusion of Neitz and Jacobs depends critically on assumptions about the relative sensitivities of the middle- and long-wave pigments near 690 nm; but it is not actually contradicted by the published molecular genetics, and there is direct microspectrophotometric evidence for individual differences in the human long-wave pigment¹³.

The bimodality reported by Neitz and Jacobs is impressive and it is difficult to imagine an experimental artefact that could generate a bimodality. But the psychophysical cognoscenti will enjoy trying to think of alternative physiological mechanisms. It might be argued that some observers do, whereas others do not, attend to a signal from a third receptor system, for the spectrum locus between Neitz and Jacobs' two primaries (546 and 690 nm) is not strictly dichromatic¹⁴. Or one might suppose that sensitivity to the 690 nm primary varies with individual differences in body temperature. In the far red, the probability of absorption of a quantum depends both on its intrinsic energy and on the thermal vibrational energy possessed by the chromophore

1. Neitz, J. & Jacobs, G.H. *323*, 623 (1986).
2. Rayleigh, Lord *Nature* **25**, 64 (1881).
3. Rushton, W.A.H., Powell, D.S. & White, K.D. *Nature* **243**, 167 (1973).
4. Rubin, M.L. *Am. J. Ophthalmol.* **52**, 166 (1961).
5. Richards, W. *J. opt. Soc. Am.* **57**, 1047 (1967).
6. Hurvich, L.M., Jameson, D. & Cohen, J.D. *Percept. Psychophys.* **4**, 65 (1968).
7. Verriest, G. *Nouv. Red. d'Optique appliquée* **1**, 107 (1970).
8. Waaler, G.H.M. *Arch. d. Norske Videnskaps-Akademi I. Mat.-Naturv. Klasse. Ny Serie* No. 9.1 (1967).
9. Waaler, G.H.M. *Arch. d. Norske Videnskaps-Akademi I. Mat.-Naturv. Klasse. Ny Serie* No. 11.1 (1968).
10. Nathans, J., Thomas, D. & Hogness, D.S. *Science* **232**, 193 (1986).
11. Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B. & Hogness, D.S. *Science* **232**, 203 (1986).
12. Mollon, J.D. *Nature News and Views* **321**, 12 (1986).
13. Dartnall, H.J.A., Bowmaker, J.K. & Mollon, J.D. *Proc. R. Soc. B220*, 115 (1983).
14. Estévez, O. thesis, Univ. Amsterdam (1979).
15. Stiles, W.S. in *Transactions of an Optical Convention marking the fiftieth anniversary of the first examination in visual optics by the Worshipful Company of Spectacle Makers*, London, 97 (1948).
16. Lewis, P.R. *J. Physiol., Lond.* **130**, 45 (1955).
17. De Vries, H. *Experientia* **4**, 357 (1948).
18. Mollon, J.D., Bowmaker, J.K. & Jacobs, G.H. *Proc. R. Soc. B222*, 373 (1984).
19. Jacobs, G.H. & Neitz, J. *Vision Res.* **25**, 141 (1985).
20. Jacobs, G.H. *Vision Res.* **24**, 1267 (1984).
21. Lyon, M. *Biol. Rev.* **47**, 1 (1972).

group of the pigment molecule^{15,16}. By manipulating his own body temperature by only one degree, the physicist De Vries was able to demonstrate changes in his visual sensitivity to the far red¹⁷. Temperature variations of this order occur diurnally; in women there is a monthly variation. If Neitz and Jacobs' young men did cluster at two temperatures, it would be easy to explain why the bimodality was obscured in their young women.

But suppose there is a polymorphism of the long-wave pigment. Has it been maintained by a heterozygous advantage? In the squirrel monkey, a basically dichromatic species, there is almost certainly a polymorphism at the single locus that specifies a pigment in the red-green spectral region^{18,19}. Heterozygous females

become behaviourally trichromatic²⁰ because X-chromosome inactivation²¹ segregates into different cones the products of the alternative alleles; the monkey's visual system seems plastic enough to exploit this added differentiation of cone cells. It has been suggested that the polymorphism is in fact maintained by the advantage to the heterozygous females¹⁸. Our own species is basically trichromatic. So if many women are heterozygous for one of the photopigments, are they in fact tetrachromatic, enjoying an extra dimension of colour discrimination? And if they are, does it give them an advantage? □

John D. Mollon is in the Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, UK.

Grand unification

Solar neutrinos may hold the key

from Lincoln Wolfenstein

THE question of the mass of the neutrino has been of great interest from Fermi's first analysis of nuclear beta decay to the present. There is still no confirmed evidence that neutrinos have a non-zero mass. However, laboratory experiments cannot probe for the very small masses that are suggested by many theories. Recently, great interest has arisen in the possibility that experiments searching for neutrinos from the Sun may prove the best way to probe neutrino mass.

A class of theories called grand unified theories treat the electron, its neutrino ν_e and the quarks that make up the nucleon as different states of a single particle. In the simplest of these theories the small neutrino mass is given by the 'see-saw' formula $m_\nu = m_D^2/M$, where m_D is the quark or electron mass and M is a mass which may be as large as the unification scale of 10^{14} to 10^{15} GeV. There are three sets, or generations, of particles. Thus in addition to the electron, there is the muon ($m_\mu = 206m_e$) and the τ particle ($m_\tau = 3,500m_e$). Correspondingly there are three neutrinos, ν_e , ν_μ and ν_τ . The theory suggests $m(\nu_e) \ll m(\nu_\mu) \ll m(\nu_\tau)$.

A method of investigating small neutrino masses is a search for neutrino oscillations. This is based on the idea that ν_e , ν_μ and ν_τ are coherent mixtures of the three neutrino mass eigenstates, in direct analogy with the quarks, where whatever causes the masses of particles mixes the eigenstates seen by the weak interaction. In the case of quarks, this mixing was first described by the Cabibbo angle. In the case of neutrinos, the electron neutrino ν_e , for example, the neutrino produced in nuclear beta decay, might be $\nu_e \cos\theta + \nu_\mu \sin\theta$, where ν_1 and ν_2 are the mass eigenstates and θ is the mixing angle. As

the neutrino propagates through the vacuum, the phase of ν_2 changes relative to ν_1 as a result of their mass difference. Thus the neutrino is no longer a pure ν_e ; elementary quantum mechanics gives the probability that it is still ν_e as an oscillating function with a minimum of $\cos^2 2\theta$. In quantum mechanics the same type of analysis describes the precession of a spin- $1/2$ magnet oriented at an angle 2θ to the magnetic field.

Many experiments at accelerators and reactors have looked for such oscillations without success. If M in the see-saw formula is greater than 10^{11} GeV the masses become very small and the distance needed for oscillations is much larger than is available in the laboratory. The theory further suggests that θ is 10° or less, so that the amplitude of vacuum oscillations is small.

Recently S.P. Mikhaeyev and A. Yu Smirnov pointed out that even with small values of θ very large oscillations can occur for neutrinos moving outward from the centre of the Sun. Their work (*Nuovo Cimento* 9C, 17; 1986) is based on my own earlier observation that neutrino oscillations in matter are modified from those in vacuum as a result of the index of refraction associated with the forward scattering of ν_e from electrons. They point out that neutrinos may pass through a region with a density such that the mixing angle in matter reaches 45° even though the vacuum θ is as small as 1° .

Over the past 20 years Ray Davis has been detecting solar neutrinos via the inverse beta-decay of ^{37}Cl . The flux observed appears to be one-quarter to one-half that expected on the basis of standard solar model calculations. This result is explained by Mikhaeyev and Smirnov

by the conversion of the missing neutrinos into ν_μ or ν_τ by the enhanced oscillation effect inside the Sun. This effect depends on the energy of the neutrinos. Davis' experiment is primarily sensitive to the high-energy ν_e from the decay of ^8B (end-point energy 14 MeV). Experiments using a gallium detector now being planned in Europe and the Soviet Union will look for the much more common low-energy neutrinos from the fundamental thermonuclear reaction $p + p \rightarrow d + \nu_e + e$ (where p is a proton, d a deuteron and e an electron). According to the oscillation theory the result of the gallium experiment depends on the neutrino mass. For a mass of the heavier neutrino (into which ν_e oscillates) of 10^{-2} eV the gallium result would agree with the standard solar model, while for lower masses the ν_e flux would be suppressed by a factor reaching 10 for a mass of $3 \cdot 10^{-4}$ eV.

If there is no suppression of these low-energy neutrinos, an alternative explanation of Davis' result is a slight reduction in the central temperature of the Sun which can cause a large decrease in the amount of ^8B in the Sun but does not affect the p - p neutrinos. The alternative can be distinguished by a measurement of the ^8B neutrinos, which would be greatly distorted in the oscillation scenario. A large liquid argon detector (called Icarus) proposed for the laboratory in the Gran Sasso tunnel in Italy by a team including Carlo Rubbia could measure this spectrum for 6 to 14 MeV. Another group propose detecting this spectrum using heavy water in a planned laboratory in a mine in Ontario, Canada.

A more dramatic way to confirm the oscillation scenario would be to detect the ν_μ or ν_τ into which the ν_e is converted. Such detection depends on the neutral current interaction and is more difficult. One possibility is to detect neutrino-electron scattering events in a detector like Icarus in an energy region where the ν_e has been found to be highly suppressed. In the heavy-water detector the neutral current interaction could result in the disintegration of the deuteron into neutron plus proton.

Measurements of the solar neutrino flux at different energies could distinguish between explanations of Davis' experiments. Whether or not the Mikhaeyev-Smirnov analysis is the correct explanation, the solar neutrino experiments can give information on neutrino mass in the range 10^{-3} to 10^{-4} eV and thus to values of M in the see-saw formula between 10^{11} and 10^{15} GeV. Thus it is hoped to obtain information on neutrino masses over a range inaccessible to experiments at accelerators and reactors. □

Lincoln Wolfenstein is Professor of Physics at Carnegie-Mellon University, Schenley Park, Pittsburgh, Pennsylvania 15213, USA.

Evolution

Contentious issues in sexual selection

from Linda Partridge and Paul Harvey

WHY do peacocks have large decorative tails? The answer is probably that females use the tails as cues in mate choice. But why do females prefer to mate with such males? We do not know despite advances in our understanding that have come from population genetic models¹⁻⁴, field experiments⁵ and breeding studies⁶. It is far from clear that research is progressing efficiently, or that any consensus on the basic issues is being reached. A recent conference on sexual selection* forced participants to confront these issues.

It could be that the elaborate and species-specific ornaments of many birds have evolved to prevent individuals from mating with members of the wrong species. Although perhaps part of the truth, that view is not fashionable. Much contemporary research is concerned with distinguishing between two other ideas. The first is that females choose mates by their ornaments because these indicate viability in other respects which will be transmitted to offspring. In other words, females get

'good' genes by mating with males that have ornaments. The second possibility is that the male ornaments impair survival, having evolved simply because females happen to prefer mating with adorned males. This was the view of Darwin who thought male ornaments required a special evolutionary explanation because they seemed detrimental to survival⁷. Two related issues must be resolved before deciding between these interpretations.

The first argument concerns fitness heritability. Until very recently, population geneticists have emphasized the theoretical prediction that populations at equilibrium under constant natural selection are not expected to exhibit heritability of total fitness^{8,9}. Taken at face value, this implies that mate choice could not result in fitter progeny. However, contemporary theory suggests that some fitness heritability can be maintained by temporally varying (particularly cycling) selection pressures or by the interaction of selection with other processes such as mutation and migration (ref. 10 and B. Charlesworth, personal communication). Such fitness heritability could be maintained in natural populations, but is it? There is general agreement that measurements of fitness heritability are urgently required, and also that the practical problems of collecting data are formidable, especially in natural populations.

The second issue concerns the assumptions, outcomes and biological relevance of particular genetic models. If a male trait is of advantage under natural selection, then female preference for it can evolve while there is genetic variability in the male trait. The reason is that females with the preference mate disproportionately with males bearing the trait, and produce fitter sons (because they bear the trait) who also carry the preference alleles which increase in frequency as a result of the association. Once such a preference reaches a sufficient frequency in the population, females gain an extra advantage by choosing males with the trait because of the attractiveness of their sons. At this point, the male ornament can become further exaggerated to a level where it becomes disadvantageous to the male's viability, Fisher's celebrated runaway process¹¹. Actually, for the runaway process to get off the ground, the male trait need not even be initially advantageous so long as some degree of female preference for it exists in the population beforehand^{2,3}. In whatever way the pro-

cess starts, at equilibrium there will be a negative genetic correlation between male mating success and male viability, in contradiction to the 'good genes' view.

A new wave of models explores the consequences of including a third character (together with the deleterious male trait and the female preference) that causes variation in male viability¹²⁻¹⁴. This can interact in various ways with the costly male trait so that, for instance, the trait is only expressed fully in the most viable males. J. Maynard Smith (Sussex University) presented a review of such models and concluded that female preference and the deleterious male trait (now an indicator of viability) can become established once mating preference alleles rise above a certain frequency.

Why are such models important and how do they relate to the notion of fitness heritability? An example is provided by Hamilton and Zuk's argument¹⁵ that bright coloration (deleterious trait) among birds may have evolved through mate choice (preference) because bright coloration is an indicator of the extent to which an individual is resistant to parasites (viability). The important point is that Hamilton and Zuk envisage that genetic variance for resistance against parasites is maintained in the population by temporally cycling selection pressures on the resistance genes of the birds. Recent genetic models of this process by Kirkpatrick and Pomiankowski demonstrate different evolutionary dynamics from those found in the basic two-character models, or the three-character models in which additive genetic variance (heritability) of viability is not maintained in the population. The relevance of such models to the real world is the subject of continued debate. If these models are important, costly sexually-selected male ornaments provide the means by which females can assess the viability of their

Sexual Selection Dahlem Conference, Berlin, FRG. 1-6 September 1986.



100 years ago

THE NEW ELEMENT, GERMANIUM

SOME months ago Dr Clemens Winkler announced the discovery of a new element which he named germanium, a preliminary account of which has already appeared in these columns. Dr Winkler has since been able to make a more systematic examination of the subject.

The following is the best method for separating the germanium. The finely-powdered mineral is intimately mixed with an equal weight of soda and sulphur, and the whole submitted to the action of a moderate red heat in a Hessian crucible. The product is powdered whilst still warm, and repeatedly boiled with water; the aqueous extract is slightly acidulated with sulphuric acid, and the precipitated sulphides of arsenic and antimony allowed to settle. On then adding a considerable excess of hydrochloric acid, the germanium sulphide is thrown down as a white voluminous precipitate; this is gently roasted, then heated with concentrated nitric acid, and finally ignited. The germanium oxide obtained may be reduced by ignition in hydrogen.

From *Nature* 34, 580; 14 October 1886.

1. O'Donald, P. *Genetic Models of Sexual Selection*. (Cambridge University Press, 1980).
2. Kirkpatrick, M. *Evolution* 36, 1 (1982).
3. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* 78, 3721 (1981).
4. Seger, J. *Evolution* 39, 1185 (1985).
5. Andersson, M. *Nature* 299 818 (1982).
6. Majerus, M.E.N., O'Donald, P., Kearns, P.W.E. & Ireland, H. *Nature* 321, 164 (1986).
7. Darwin, C. *The Descent of Man and Selection in Relation to Sex*. (Murray, London, 1871).
8. Williams, G.C. *Sex and Evolution* (Princeton University Press, 1975).
9. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1976).
10. Partridge, L. in *Mate Choice* (ed. Bateson, P.) 227 (Cambridge University Press, 1983).
11. Fisher, R.A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
12. Andersson, M. *Biol. J. Linn. Soc.* 17, 375 (1982).
13. Dominey, W.J. *J. theor. Biol.* 101, 495 (1983).
14. Andersson, M. *Evolution* 40, 804 (1986).
15. Hamilton, W.D. & Zuk, M. *Science* 218, 384 (1982).
16. Huxley, J. *Evolution, the Modern Synthesis* (Allen & Unwin, London, 1942).
17. Dobzhansky, Th. *Genetics and the Origin of Species* 3rd edn. (Columbia University Press, 1951).
18. Mayr, E. *Systematics and the Origin of Species* (Columbia University Press, 1942).
19. Barton, N.H. & Hewitt, G.M. *A. rev. Ecol. Syst.* 16, 113 (1985).
20. Levin, D.A. *Evolution* 39, 1275 (1985).
21. Armbruster, W.S. *Evolution* 39, 733 (1985).

mates, and hence of their offspring.

Empirical investigation of this issue is both desirable and difficult. All 'good genes' models would be falsified if the assumption of heritability of fitness were violated. A prediction of some models would be disproved if there was not a positive genetic correlation between male mating success through mate choice and the viability of female progeny. Both of these points could be investigated experimentally. An alternative way forward may be by using the comparative method. The scientific world took Hamilton and Zuk's ideas more seriously because the authors were able to show a correlation between parasite loads and bright coloration across passerine birds.

It became obvious during the meeting that we would like to know more about female mating preferences. Many of the theoretical models are very sensitive to violation of the assumption that selection does not act directly on the mating preferences themselves. In more biological terms, if there are direct benefits or costs of choice to the female (in terms of her own viability or fertility), much of the current theory may not be relevant, and a restricted range of evolutionary outcomes is expected. Yet mate choice often results in direct benefits, such as improved territories. There may be ubiquitous costs of choice in terms of time wasted or increased susceptibility to predation. Empirical work on this issue is essential. On the other hand, theory needs to be developed to understand how the joint evolution of female preferences for resources provided by males and for male traits proceeds.

Another important but virtually unstudied aspect of mate preferences is that of the rules females use to choose their mates, yet the outcome of genetic models for the evolution of deleterious male traits can differ markedly depending on the rule used⁴. Do females inspect several males and then choose one, or do they have a fixed probability of accepting a male with a particular phenotype? Are preferences open-ended, so that males with, say, longer tails are always preferred, or is there an upper limit on the preference? And do females of different species have different rules for choice? Answering these questions could keep many of the participants of the meeting busy for years to come.

Finally, we know very little about the extent to which mating preferences have evolved in contexts outside sexual selection. Perceptual systems may be biased in certain directions because of, for instance, selection for detection of food or predators. If this is so, the initial evolution of female mating preferences may require no special explanation. Burley's reported demonstration that female zebra finches (*Poephila guttata*) prefer males wearing large white hats could be an example of

such a predisposition.

A fascinating sociological change has occurred since the Modern Synthesis¹⁶⁻¹⁸. Darwin noted that related species often differ most in male secondary sexual characters. The architects of the Modern Synthesis, being particularly interested in speciation, considered that such differences had arisen as a result of selection for reproductive isolation. Despite little change in the data, recent discussions have emphasized the importance of sexual selection acting independently in genetically isolated populations. Part of the reason for this shift in attitudes may have come from theoretical work showing that

the conditions for the evolution of pre-mating isolation in hybrid zones are restrictive¹⁹. Cases of pre-mating isolation between species that have never met cannot be explained by selection for reproductive isolation, but the existence of such examples does not mean that this type of selection is unimportant. In some cases pre-mating isolation is stronger in areas of overlap between closely related species than in areas where they occur separately^{20,21}. The shift in emphasis from species isolation to independent sexual selection is probably unjustified.

Despite our concern that species-isolating mechanisms have gone out of fashion as an explanation for the evolution of bizarre male characters, the relative roles of the runaway process and good genes models make fascinating topics for research. Rarely does one return from a conference knowing that the important controversial issues were fully discussed and that proponents of alternative views more clearly understood each other's position. The genetic consequences for the population, the behavioural rules, the perceptual predispositions involved, the costs and benefits incurred and the genetic basis of mate choice are topics that we must know more about if we are to understand fully the evolution of secondary sexual characters. And that is what sexual selection is all about. □

Linda Partridge is in the Department of Zoology, University of Edinburgh, Edinburgh EH9 3JT and Paul Harvey is in the Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

Count Raggi's bird of paradise (Frank Lane).

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Plate tectonics

Response to an Archaean continental collision

from Mike Bickle

SEDIMENTS provide some of the most comprehensive records of tectonic processes in the Earth's history. Therefore, the recent report by K. Burke, W.S.F. Kidd and T.M. Kusky (*Tectonics* 5, 439; 1986), which shows that the late Archaean Witwatersrand Basin, South Africa was deposited as a foreland basin in response to processes at a convergent plate boundary, is of considerable importance.

Direct evidence for Archaean orogenic belts analogous to modern plate-tectonic-related orogenic belts has proved surprisingly elusive despite the popularity of the notion that plate tectonics was operating as long ago as the Archaean. The large area and state of preservation of the Witwatersrand sequence and the vast amount of geological information available from the intense gold-mining activity

in the Witwatersrand Basin should provide a detailed record of the tectonic evolution of the basin and its source regions, and these data have the potential to produce an unrivalled study of the tectonic cause and setting of this Archaean sedimentary sequence.

Burke *et al.* suggest that the Witwatersrand foreland basin developed on the Kaapvaal Craton in response to loading of the crust by orogenic activity to the north. They envisage the development of an Andean-type arc on the northern margin of the Kaapvaal Craton terminated by the subsequent collision of the Zimbabwe and Kaapvaal cratons about 2,700 million years ago. The Limpopo granulite facies province may represent the site of a Tibetan style collision-related plateau. Several characteristics of the Witwatersrand

Basin support this hypothesis: it is asymmetrical, both in thickness and in deformation style, compatible with a major mountain range to the north. Although there is relatively little magmatism, the sediments contain silicic volcanic detritus. There is progressive migration of the depositional axis and the basin shrank in area over a timescale consistent with the expected visco-elastic response of a loaded lithosphere.

The interpretation of the Witwatersrand Basin as a foreland basin provides a rather better explanation for some of its features than the earlier idea that it developed in response to crustal stretching. The basin lacks the well-developed initial active-rift-related sedimentary and magmatic phase that is expected to result from crustal stretching. The upper part of the Witwatersrand sequence, the Central Rand Group, is dominated by molasse-type sediments deposited as coalesced fluvial-alluvial fans in response to increasing proximity of tectonic activity. The higher stratigraphic units in stretching-related basins exhibit progressively slower subsidence and sedimentation. One of the other very few similar major deposits of this age, the Fortescue Group in the Pilbara Block, Western Australia, has been interpreted as deposited on rifted crust (Blake, T.S. *Nature* **307**, 721; 1984).

The foreland-basin interpretation has some problems. Burke *et al.* argue that the igneous activity, mainly restricted to the lowermost protobasinal Dominion Group, is of Andean arc-type. A recent description of these volcanics (Bowen, T.B. *et al. Precambrian Res.* **31**, 297; 1986) describes the volcanics as a bimodal tholeiitic-basalt-rhyolite association, consistent with a rift-related setting rather than an Andean arc setting. A.J. Tankard *et al. (Crustal Evolution of South Africa Springer, 1982)* indicate entry points for sediments into the basin from the south-west and east as well as the dominant supply from the north. Both the igneous association and the geography of the source terrain could reflect the geological complexities expected in a large area affected by a major continental collision.

The geochronology of the Witwatersrand sequence frustrates attempts at tectonic synthesis but ultimately should provide crucial information about the genesis of the basin. The age of the sequence is thought to be between 2,800 and 2,300 million years. It is possible, although unlikely, that the Witwatersrand sequence comprises separate sedimentary sequences deposited in response to two or more unrelated tectonic events spaced over hundreds of millions of years. The poor constraints on the timing reduce substantially the confidence that can be placed on the attempts of Burke *et al.* to correlate the formation of the Wit-

watersrand Basin with geological events in the Kaapvaal Craton and Limpopo Province to the north. The problem should not be insoluble. High-resolution ion-microprobe uranium-lead zircon analyses combined with high-precision uranium-lead zircon ages should allow resolution of the age and duration of the Witwatersrand sequence to within a few million years. Further, ion-microprobe uranium-lead

analyses of detrital zircons and zircons from rock fragments in the Witwatersrand sequence could reveal much of the geological history of the source terrains. Until the chronological studies are satisfactory the tectonic interpretations will remain uncertain. □

Mike Bickle is in the Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK.

Cancer

Malignant tumours generated by recessive mutations

from Henry Harris

AS FAR as I know, the idea that malignant tumours might arise as a consequence of recessive mutations in somatic cells first emerged from our cell fusion experiments in which malignancy was found to be suppressed when malignant cells were crossed with non-malignant ones¹. In 1971 Ohno² argued, on theoretical grounds, that malignancy was much more likely to be generated by recessive mutations in somatic cells than by dominant ones and suggested that the recessive mutations might be unmasked by genetic events that resulted in loss, or effective loss, of the unaffected gene on the homologous chromosome. This state of hemizyosity could commonly arise from the loss of the homologous chromosome itself. In the same year a specific version of this general model was proposed by Knudson³ to account for the incidence pattern of the malignant eye tumour, retinoblastoma. On page 643 of this issue⁴, Friend *et al.* come close to identifying the gene that is implicated in retinoblastoma.

Retinoblastoma occurs in two forms, one sporadic and the other in which the predisposition to form the tumour is inherited (in an autosomally dominant fashion). Knudson postulated that the genetic loss involved was the same in the two forms of the disease, but that in the heritable form one of the alleles was already mutated or otherwise rendered non-functional in the germ line. A single event inactivating the homologous locus in the unaffected chromosome would then be enough to generate the tumour in the predisposed individual; in sporadic cases both alleles would have to have been inactivated by two separate genetic events. This model, involving homozygosity or hemizyosity of recessive mutations, accommodated the incidence data reasonably well and prompted a search for more direct supporting evidence. This was in due course obtained.

It is now clear, from cytogenetic observations, electrophoretic examination of a

linked enzyme polymorphism and analysis of DNA restriction fragment length polymorphisms⁵⁻⁷ that Knudson's model is essentially correct. The locus involved maps to band q14 on the long arm of chromosome 13 and has been given the name *Rb*. In the cases where predisposition is heritable, the somatic cells of the affected individual often show a visible deletion in this chromosome region; and where no deletion is visible, other forms of analysis suggest the presence of a sub-microscopic deletion. Tumours arising in such individuals commonly show either homozygosity of the mutated allele or hemizyosity caused either by a secondary deletion in the previously unaffected allele or by the elimination of the whole of the unaffected chromosome 13.

In the paper by Friend *et al.*⁴ the genetic investigation of retinoblastoma is taken a step further. By the use of standard recombinant DNA technology, these authors have located a coding sequence at the q14 region of chromosome 13 that might well be the *Rb* locus. Essentially two lines of evidence are presented in support of this view. The first is that the complementary DNA probe (p4.7R) that recognizes this DNA sequence fails to reveal any corresponding RNA transcripts in four retinoblastoma cell lines and one osteosarcoma cell line. (Osteosarcomas are common secondary tumours in patients with the heritable form of retinoblastoma.) But such RNA transcripts were found in one human retinal cell line transformed by adenovirus 12 and in a range of other malignant tumours. The second line

1. Harris, H. *et al. Nature* **223**, 643 (1969).

2. Ohno, S. *Physiol. Rev.* **51**, 496 (1971).

3. Knudson, A.G. *Proc. natn. Acad. Sci. U.S.A.* **68**, 820 (1971).

4. Friend, S.H. *et al. Nature* **323**, 615 (1986).

5. Sparkes, R.S. *et al. Science* **219**, 971 (1983).

6. Benedict, W.F. *et al. Science* **219**, 973 (1983).

7. Cavenee, W.K. *et al. Science* **228**, 501 (1985).

8. Harris, H. *J. Cell Sci. Suppl.* **4**, 431 (1986).

9. Sager, R. *Adv. Cancer Res.* **44**, 43 (1985).

10. Capon, D.J. *et al. Nature* **304**, 507 (1983).

11. Santos, E. *et al. Science* **223**, 661 (1984).

12. Adams, J.M. *et al. Nature* **318**, 533 (1985).

of evidence is that some 30 per cent of a collection of 40 retinoblastomas, 8 osteosarcomas and 2 undifferentiated tumours arising in retinoblastoma patients showed some structural abnormality in their DNA when fragments generated by restriction enzymes were examined by the p4.7R probe.

The authors do not claim that the evidence is decisive. The failure to detect any alteration in 70 per cent of cases may, as the authors suggest, be explained by the presence of genetic lesions, such as point mutations, that are not detected by the examination of fragments generated by restriction enzymes.

Alternatively, it is possible that the locus detected by the p4.7R probe is not the *Rb* locus itself, but a closely linked locus involved in more complex chromosome rearrangements. This possibility is perhaps made less likely, but is not eliminated, by the fact that in one retinoblastoma and one osteosarcoma, internal deletions were found within the coding region identified by the p4.7R probe. Friend *et al.* contend that this probe should be a powerful diagnostic reagent for distinguishing between the heritable and the sporadic form of retinoblastoma; but it is not immediately clear in what way a probe that yields false negatives in 70 per cent of cases is likely to be particularly useful in clinical diagnosis.

In any case, it should not be long before the identity of the genetic region detected by this probe is firmly established. If this region is the *Rb* locus itself, we can expect soon to receive some sequence information that might provide a clue to the function of the protein that it specifies. Because I adhere to the view that the progressive multiplication of cancer cells is a secondary consequence of their having sustained some genetic alteration that prevents them from completing their normal programme of differentiation, my guess is that the *Rb* locus will be found to specify a protein that is essential for the execution of that programme in the retinal cell.

There is good evidence that homozygosity or hemizyosity of recessive mutations is responsible for tumorigenesis in other malignancies in which predisposition to form the tumour is inherited in a mendelian fashion, for example Wilms' tumour, hepatoblastoma, rhabdomyosarcoma and familial renal carcinoma. But Knudson has argued that there may be a cryptic genetic predisposition in many other tumours, including the common malignancies. In that case, recessive mutations might be responsible for generating a much wider range of malignant tumours than was originally suspected. Certainly the great majority of tumour cells, irrespective of the cell type or the aetiology of the malignancy, and even when the cells bear a well-defined oncogene that is actively transcribed, yield

First glimpse of the *DMD* gene?

A DECADE ago, molecular analysis of the gene responsible for the X-linked disease Duchenne muscular dystrophy (DMD) seemed an unreal dream. Five years ago, the identification of restriction fragment length polymorphisms genetically linked to the *DMD* gene¹ began to turn the dream into reality. On page 646 of this issue², Louis Kunkel, Anthony Monaco and colleagues make further progress with the report of partial characterization of transcripts that probably derive from the *DMD* gene.

Although the results of Monaco *et al.* represent the culmination of a series of brilliant experiments from one laboratory, these experiments would not have been possible without the contribution of many others. Clinicians have provided material from patients (a recent paper³ acknowledges donations from 25 clinical laboratories); cytogeneticists have defined chromosomal translocations and deletions; somatic cell geneticists have isolated important chromosomes on rodent cell backgrounds; and molecular geneticists have provided a wealth of cloned probes for use as markers in family studies.

These combined efforts resulted in the precise definition of the chromosomal location of the *DMD* gene³. As I discussed recently in *News and Views*⁴, knowledge of the chromosomal location and the availability of an X chromosome deleted for the *DMD* gene were exploited by Kunkel and colleagues to isolate sequences from the immediate vicinity of the *DMD* locus⁵.

A slightly different approach was taken by Ray *et al.*⁶ who cloned the sequences present at a chromosomal translocation breakpoint which was associated with *DMD*. The new genetic analysis⁷ using both of these sequences suggests that the *DMD* gene is very large, very complex or both. During these studies Monaco *et al.* isolated more than 200 kilobases of DNA from the vicinity of the *DMD* gene. These clones were instrumental for isolating the gene for chronic granulomatous disease⁸ and have

now been screened for expressed sequences related to *DMD*.

Rather than screen valuable messenger RNA with a large number of unique sequence probes Monaco *et al.* first searched for cross-species DNA homology in Southern blots. They argued that any expressed exons would show evolutionary conservation, and found two candidate probes that hybridized to DNA from all mammals tested.

One of these probes, PERT-25, reacts with a 16-kilobase transcript which is present in fetal muscle RNA samples but is not detectable with messenger RNA from cultured human myoblasts or HeLa cells. Short complementary DNA clones, isolated after screening a fetal muscle complementary DNA library with PERT-25, recognize exons distributed more than 110 kilobases of DNA in the *DMD* region. By extrapolation to full-length clones this candidate *DMD* gene may be 1,000–2,000 kilobases long.

As with all scientific advances these results pose many questions which will be difficult to answer. Obtaining the full-length sequence of a mammalian 16-kilobase messenger RNA has only been achieved for one other transcript (apolipoprotein B⁸⁹) and the possibility that this candidate *DMD* gene produces multiple transcripts with alternative exon usage will also keep many molecular biologists employed. Finally, obtaining the formal proof that this is the *DMD* gene may be the hardest of all. Nevertheless, the new work is good news to families afflicted with Duchenne muscular dystrophy.

Peter N. Goodfellow

1. Murray, J.M. *et al.* *Nature* **300**, 69 (1982).

2. Monaco, A.P. *et al.* *Nature* **323**, 646 (1986).

3. Kunkel, L. *et al.* *Nature* **322**, 73 (1986).

4. Goodfellow, P. *Nature News and Views* **322**, 12 (1986).

5. Kunkel, L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 4778 (1985).

6. Ray, P.N. *et al.* *Nature* **318**, 672 (1985).

7. Royer-Pokova, B. *et al.* *Nature* **322**, 32 (1986).

8. Knott, T.J. *et al.* *Nature* (in the press).

9. Yang, C.-Y. *et al.* *Nature* (in the press).

non-malignant hybrids when they are fused with normal diploid fibroblasts⁸.

Are there any cellular genes that act to produce tumours in a genetically dominant fashion? They have yet to be found. As several authors have pointed out (see refs 8 and 9) the fact that certain cellular genes when transfected into mouse fibroblast NIH 3T3 cells can induce morphological transformations in these cells and eventually render them tumorigenic, is no evidence that the transfected genes (oncogenes) act in a genetically dominant fashion. The changes that this procedure produces in the genome of the NIH 3T3 cell are very complex and preclude interpretation in formal genetic terms. Where the question has been examined in nat-

urally occurring malignant epithelial tumours^{10,11}, it has been found that the presence of a mutated oncogene in the tumour is often associated with the absence of its normal allele, a situation very reminiscent of retinoblastoma. In the case of the characteristic translocations involving the *c-myc* oncogene in lymphoid neoplasms, experiments with transgenic mice again indicate that, in addition to the translocation, a second genetic event must occur in the cell before it can generate a malignant tumour¹². □

Henry Harris is Regius Professor of Medicine and head of the Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK.

Anaesthesia

Models of consciousness

from Keith W. Miller

How simple molecules cause general anaesthesia has been a puzzle for 140 years. It is agreed that the potency of the structurally diverse general anaesthetics is proportional to their solubility in lipid bilayers and that pressure can reverse general anaesthesia. To some this suggests that a single molecular mechanism giving rise to bilayer expansion or disordering is involved. To others the same observations implicate protein-anaesthetic interactions, probably involving separate sites for each of the major structural classes of anaesthetic. Ignorance of the physiology of consciousness forces these seekers of the molecular mechanisms of general anaesthesia to study simple but arbitrarily chosen model systems, such as lipid bilayers or pure proteins, in the hope that eventually the behaviour of the whole may be reconstructed from the sum of its parts. Few have been bold enough to attempt such a synthesis, but recently Alec Bangham and Martyn Hill¹, building on previous work on their proton pump/leak hypothesis, have made such an attempt.

Bangham and Hill set out to explain how consciousness can be lost not just by the action of an anaesthetic, but also by other perturbations such as cold or anoxia. Neuronal membranes are involved in maintaining a non-equilibrium balancing act in which the tendency of ion concentration gradients to decay is balanced by energy-requiring pumps. In such a steady-state system any change in the opposing fluxes will alter the point of equilibrium. The pumps are both energy and temperature dependent, but little affected by general anaesthetics, whereas passive ion permeability (leak) is known to be increased by general anaesthetics. If it is assumed that anaesthesia results in some way from this increased leak discharging the ion gradients, then cold would have the same net effect by slowing the energy-dependent pumping of ions, hence inducing cold insensibility. Similarly one can explain other changes in the conscious state, such as the fact that pressure reverses the leak and thus anaesthesia; or that anoxia interrupts energy supplies to the pumps and leads to unconsciousness. One specific example of these effects is that the ATP-dependent accumulation of the neurotransmitter dopamine

in rat brain synaptic storage vesicles is reversed by general anaesthetics, perhaps because an anaesthetic-induced inward leak of cations facilitates the outward leak of hydrogen ions, increasing the proportion of dopamine in the uncharged, membrane-permeable form.

The hypothesis can be criticized on several grounds. For example, most ion pumps are unlikely to be overwhelmed by leaks of the magnitude induced by anaesthetics, although such a situation could

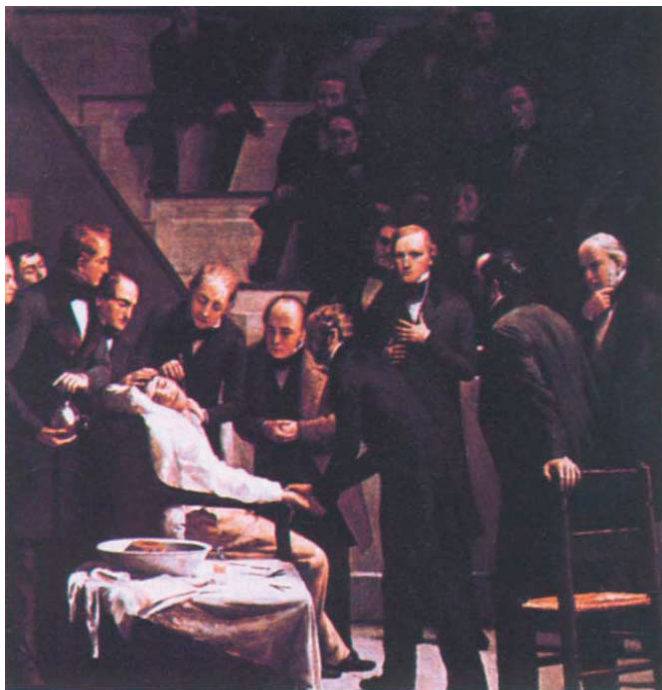
Specific effects of general anaesthetics are well known — the only controversy concerns whether they cause or accompany anaesthesia (see refs 4–7 and references cited therein). Allosteric actions of barbiturates have been demonstrated at the voltage-sensitive sodium channel as well as at the γ -aminobutyric acid and nicotinic acetylcholine receptors. These actions on the receptors are stereoselective, and saturable binding of radiolabelled barbiturates to the acetylcholine receptor can be demonstrated. However, the pharmacology of barbiturates suggests these actions are less related to general anaesthesia than to anticonvulsant or other actions.

On the other hand, gaseous anaesthetics act on channels by a mechanism independent of these allosteric sites. In the voltage-sensitive sodium channel and in the acetylcholine receptor these anaesthetics reduce the number of channels in the resting state. In both channels this stabilization of the inactivated or desensitized state is pressure-reversible. How common are such actions and what structural features underlie them? More systematic information might provide an indication of the physiological sites involved.

Traditional studies on lipid bilayers have not pointed to such sites because there are membranes of appropriate composition in most cells. This situation could be about to change as a result of intriguing new evidence from *in vivo* studies^{7,8} of adaptive changes in membrane lipid composition in animals chronically exposed to ethanol. Although the adaptive change in average physical properties is disturbingly small, there is growing evidence that it is confined to specific lipid species, implying a much larger change in some microregion of the membrane⁸. If this proves to be the case, it would be a useful pointer to the physiological site of action of ethanol and perhaps of other general anaesthetics. Without some such pointers the way may be long. □

1. Bangham, A.D. & Hill, M.W. *Chem. Phys. Lipids* **40**, 189 (1986).
2. Adams, P.R. & Galvan, M. *Adv. Neurol.* **44**, 137 (1986).
3. Roth, S.H. *Fdn Proc.* **39**, 1595 (1980).
4. Franks, N.P. & Lieb, W.R. *Nature* **300**, 487 (1982).
5. Miller, K.W. *Int. Rev. Neurobiol.* **27**, 1 (1985).
6. Roth, S.H. & Miller, K.W. (eds) *Molecular and Cellular Mechanisms of Anesthetics* (Plenum, New York, 1986).
7. Alcohol and the Cell *Ann. N.Y. Acad. Sci.* (in the press).
8. Tarachi, T.F., Ellington, J.S., Wu, A., Zimmerman, R. & Rubin, E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Keith W. Miller is Edward Mallinckrodt Professor of Pharmacology in the Department of Anaesthesia, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.



The first public demonstration of surgery under general anaesthesia on 16 October 1846, exactly 140 years ago, in the Bulfinch Building, now known as the Ether Dome at the Massachusetts General Hospital. The painting is by Robert Hinckley.

arise locally if the pump density were low. Nonetheless, its emphasis on the need to consider the inter-relationship between separate components of a neurone is valuable. This emphasis is echoed by recent attempts to explain the behaviour of intact neurones in terms of the plethora of ion channels that have been discovered in the past decade. These channels are often coupled to each other by changes in voltage in ways that may profoundly modify overall electrical activity. This modulation in turn may be regulated by the activities of neurotransmitters and other agents acting on allosteric sites (see ref. 2 for a review). That general anaesthetics can modulate such channels is suggested by the drastic modification that low concentrations of halothane induce in the stable rhythmical discharge exhibited by a crayfish stretch receptor under tension³.

How do bacteria acquire plant genes?

SIR—Carlson and Chelm¹ have presented evidence of gene transfer from a plant to a closely associated bacterial symbiont of the family Rhizobiaceae. But one particular feature of the purportedly transferred *glnI* gene, namely the lack of introns, suggests an alternative mechanism by which the gene transfer occurred.

There is good reason to suppose that introns in eukaryotic genes have an ancient origin^{2,3}, so it seems unlikely that an intronless form of the gene was transferred. The authors propose that the recipient Rhizobiaceae acquired DNA encoding a homologue of the cellular glutamine synthetase (GS) gene, complete with introns, which were subsequently deleted in the absence of any obvious selective advantage to cells possessing two GS genes. However, the alternative explanation that immediately presents itself is that the Rhizobiaceae have acquired a processed gene or pseudogene^{4,5} of plant GS derived from transcribed RNA. Mediation by a plant virus⁶ can even be considered.

The incorporation of processed genes or pseudogenes into the eukaryotic genome is well documented. Reverse transcriptase-like activity "may be the major cause of movement of mobile DNA elements within the nuclear DNA of perhaps all eukaryotes"² and, given the evidence of Carlson and Chelm¹, perhaps also between eukaryotes and prokaryotes.

A.G. BROWNLEE

CSIRO,
Division of Animal Production,
PO Box 239,
Blacktown, NSW 2148, Australia

1. Carlson, T.A. & Chelm, B.K. *Nature* **322**, 568–570 (1986).
2. Darnell, J.E. & Doolittle, W.F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1271–1275 (1986).
3. Senapathy, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2133–2137 (1986).
4. Sharp, P.A. *Nature* **301**, 471–472 (1983).
5. Baltimore, D. *Cell* **40**, 481–482 (1985).
6. Pfeiffer, P. & Hohn, T. *Cell* **33**, 781–789 (1983).

Caught in a trap

SIR—We were pleased to read your report¹ "Catching atoms in beams of light" on our recent experiment². It contains, however, an unfortunate misconception: specifically, that the essential requirement for success was the suspension of conventional belief as "outlined" in the paper of Pritchard *et al.*³ on the scattering force and the optical Earnshaw theorem. Since, as your report correctly states, our trap is due to the dipole force, the work of Pritchard *et al.*, which contains new trapping proposals based on the scattering or spontaneous force, does not apply to our trap and had no influence on our experimental work. In fact, one of the principal conclusions of our paper is that the trapping

results are in excellent agreement with previous existing theoretical and experimental knowledge as developed over the past 16 years, and thus are in no way contrary to "conventional belief".

Pritchard's views of the optical Earnshaw theorem⁴ and scattering force traps, as quoted in your report, involve issues that are quite separate from the above. We feel that his opinion, stating in essence that the optical Earnshaw has somehow been an obstacle to valid laser trapping proposals, is not supported by the published literature.

A. ASHKIN
J.E. BJORKHOLM
STEVEN CHU

Electronics Research Laboratory,
AT&T Bell Laboratories,
Holmdel, New Jersey 07733, USA

1. *Nature* **322**, 403 (1986).
2. Steven Chu, Bjorkholm, J.E., Ashkin, A. & Cable, A. *Phys. Rev. Lett.* **57**, 317 (1986).
3. Pritchard, D.E., Raab, E.L., Bagnato, V., Wieman, C.E. & Watts, R.N. *Phys. Rev. Lett.* **57**, 310 (1986).
4. Ashkin, A. & Gordon, J.P. *Opt. Lett.* **8**, 511 (1983).

Cancer risk assessments in light of Chernobyl

SIR—The radiation doses received by the population of Europe as a result of the Chernobyl fall-out have stimulated a reappraisal of the biological effects of ionizing radiation at low levels of exposure. Although the International Commission on Radiological Protection (ICRP) argue that an assumption of linearity at low dose is 'over-cautious', there is evidence from animal experimentation² and human observation³ that high LET (linear energy transfer) radiation becomes more dangerous per unit dose at lower doses and not less dangerous. Whereas these observations are of particular relevance to the excess cases of child leukaemia observed around nuclear facilities in the United Kingdom^{4–7} they also emphasize that ICRP is not as cautious as it might appear. Furthermore, it is difficult to understand why ICRP waited until 1957 before adopting a non-threshold model. Ten years previously, D.G. Catcheside submitted a paper to the MRC Radiation Protection Subcommittee which stated: "All quantitative experiments show that even the smallest dosage of radiation produces a genetic effect, there being no threshold dose below which no genetic effect is induced"⁸.

The ICRP risk estimates are based mainly on atomic bomb data, even though the dosimetry is known to be unreliable and is now undergoing extensive revision⁹. Table I includes data from a paper by Charles and Lindop¹⁰, which analyses biological data in UNSCEAR 1977¹¹ but excludes the atomic bomb studies. The ICRP estimates 125 fatal cancers per rem per million persons exposed, which is less than one-third of the upper estimate offered by UNSCEAR, and one-quarter of the upper estimate offered by the US National Academy of Sciences BEIR III Committee¹². The risks are more than doubled if cancer incidence, rather than cancer mortality is considered, producing ranges of 260–880 for males, and 550–1,620 for females¹³. Because public exposure limits around nuclear facilities in the United Kingdom at 0.5 rem per annum are twenty times higher than those allowed in the United States, Germany or Japan, it is difficult for governments relying on the ICRP to claim that their standards of safety are amongst the highest in the world.

Furthermore, the recent MRC mortality study of the UK nuclear workforce demonstrates dose-dependent relationships for exposure to ionizing radiation and excess cases of both cancer and leukaemia three times greater than those predicted by the ICRP "assuming that these risks are spread over a twenty-five year period"¹⁴. In fact cancer risk estimates are spread over a forty year period. Therefore the observed increases are five times greater than those predicted by ICRP. On this basis, only deep sea fishing is more dangerous than working in the nuclear industry.

These considerations are also of relevance when calculating the biological impact of the Chernobyl fall-out on the UK population. The population of North Wales, Northern Ireland, Scotland and North-West England will receive 400,000 rems over the next five years¹⁵. Using a linear non-threshold model, one can expect 50 extra cancer fatalities according to ICRP, 50–176 according to UNSCEAR, and up to 500 fatal and non-fatal cancers (excluding skin cancer), according to BEIR III. When calculating effects in other parts of Europe, it is essential that these different estimates are taken into consideration. It is also apparent that for the second time in its short history, there are major discrepancies between the re-

Table I Life-time cancer risk per rem per million persons exposed (assuming linearity)

Authority	Year	Fatal cancers (including leukaemia)	All cancers (excluding skin)
ICRP	1977	125	
UNSCEAR*	1977	100–440	300–700
BEIR III	1980	167–501	260–880 (males) 550–1,620 (females)

*Excludes the atomic bomb studies.

commendations of the ICRP and the scientific data.

ROBIN RUSSELL JONES
FRIENDS OF THE EARTH POLLUTION
ADVISORY COMMITTEE,

c/o The Old Cottage
Waxham Street
Stoke Poges
Bucks SL3 6NB, UK

1. *Recommendations of the International Commission on Radiological Protection* (ICRP Publ. No. 26) (Pergamon, Oxford, 1977).
2. Muller, W.A., Gossner, W., Hug, O. & Luz, A. *Health Phys.* **35**, 33–55 (1978).
3. Mays, C.W., Spiess, H. & Gerspach, A. *Health Phys.* **35**, 83–90 (1978).
4. *Investigation of the Possible Incidence of Cancer in West Cumbria: Report of the Independent Advisory Group* (HMSO, London, 1984).
5. Heasman, M.A., Kemp, I.W., Urguhart, J.D. & Black, R. *Lancet* **i**, 266 (1986).
6. Heasman, M.A. *et al.* *Lancet* **i**, 1188–1189 (1984).
7. Barton, C.J., Roman, E., Ryder, H.M. & Watson, A. *Lancet* **ii**, 1248–1249 (1985).
8. Catcheside, D.G. *Genetic Effects of Irradiation with Reference to Man: MRC Committee on the Medical and Biological Applications of Nuclear Physics Protection Subcommittee* MRC Publ. No. 47/77 (1977).
9. Loewe, E.W. & Mendelsohn, E. rep. D80–14. Lawrence Livermore National Laboratory [Available from National Technical Information Service, Springfield, Va 22216].
10. Charles, M.W. & Lindop, P.J. *J. Soc. for Radiological Protection* **i**, 15–19 (1981).
11. United Nations Scientific Committee on the Effects of Atomic Radiation, *Sources and Effects of Ionising Radiation* (United Nations, New York, 1977).
12. Beir III Committee, *The Effects on Populations of Exposures to Low Levels of Ionising Radiation* (The National Academy of Sciences, Washington, 1980).
13. Radford, E.P. *Radiat. Res.* **84**, 369–394 (1980).
14. Beral, V. *et al.* *Br. med. J.* **291**, 440–447 (1985).
15. Fry, F.A., Clarke, R.H. & O'Riordan, M.C. *Nature* **321**, 193–195 (1986).

Declining support for Imanishi

SIR—Last year, Halstead criticized Imanishi's evolutionary theory saying that "Imanishi is not simply a popular writer. In Japan he is often held up as an intellectual giant equivalent to Charles Darwin"¹. Subsequently, two letters from Japanese scientists, Sibatani² and Nakahara, *et al.*³, seem to praise Imanishi and his evolutionary theory. We fear that readers will receive the erroneous impression that many Japanese scientists oppose darwinian theory and agree with Imanishi's theory, whereas the truth is that the influence of Imanishi and his theory are declining in Japan.

Imanishi greatly influenced Japanese primatologists at one time. Under his supervision, many creative studies were carried out on the social systems of Japanese macaques in the 1950s and 1960s. But times have changed. From 1983 to 1986, a large-scale project on optimal strategy and social structure was sponsored by the Ministry of Education, Science and Culture. Most of the ecologists working within it did so within a darwinian framework. And at the international symposium marking the end of the project, only a few scientists seemed to disagree with darwinian theory and support Imanishi.

Recently, Kishi⁴ and Kawata^{5,6} have pointed out that Imanishi's theory slowed

the development of darwinian theory in Japan. The "Imanishi school" is still active, especially in primate ecology and sociology. Although Japanese primatologists, including us, are still hampered by the influence of Imanishiism, we criticize his theory and try to eliminate its negative influence^{7,8}.

We argue that Imanishi's description of phenomena is not incompatible with darwinian theory that deals with mechanisms of evolution. Although Imanishi himself considers his theory to be "anti-Darwinism", his view is completely misleading^{5–7}; his theory is but a part of darwinian theory. We think that Imanishi's theory should be reconstructed within a darwinian framework, thereby contributing to the development of theories of mechanisms of evolution^{7,8}. Undoubtedly it is nonsense simply to say, "Darwinian theory is bad", or, "Imanishi's theory is nonsense". To develop scientific theories, it is important to consider and discuss theories and opinions without dogmatism.

OSAMU SAKURA
TOSHIYUKI SAWAGUCHI
HIROKO KUDO
SHIN'ICHI YOSHIKUBO

Primate Research Institute,
Kyoto University,
Inuyama 484, Aichi, Japan

1. Halstead, B. *Nature* **317**, 587–589 (1985).
2. Sibatani, A. *Nature* **320**, 492 (1986).
3. Nakahara, H., Sagawa, T. & Fuke, T. *Nature* **321**, 475 (1986).
4. Kishi, Y. *Seibutsu Kagaku (Biological Science)* **38**, 104–110 (1986).
5. Kawata, M. *Seibutsu Kagaku* **37**, 72–78 (1985).
6. Kawata, M. *Seibutsu Kagaku* **38**, 31–37 (1986).
7. Sakura, O. *Seibutsu Kagaku* (in the press).
8. Sawaguchi, T. *Seibutsu Kagaku* **38**, 160–166 (1986).

Glucagon receptor number and the MHC

SIR—The recent report of glucagon activation of both cyclic AMP and inositol phosphate pathways by Wakelam *et al.* was interpreted in terms of two different glucagon receptors¹. Alternatively, there could be one glucagon receptor molecule with differing association constants for glucagon (or its analogue, TH-glucagon¹) determined by regulatory proteins which also direct whether adenylate cyclase or inositol phosphate breakdown will be stimulated. The genetic effects of the major histocompatibility complex (MHC) class I antigens on both adenylate cyclase and membrane phospholipid pathways seems more compatible with the hypothesis of one receptor protein.

Originally, the mouse MHC (H-2) was found to influence liver cAMP levels and glucagon binding to hepatocyte membranes^{2,3}. More recently, the genomic region encoding H-2 has also been found to influence membrane methyltransferase I, which synthesizes phosphatidylmonomethylethanolamine from phosphatidylethanolamine⁴. This methylation con-

sumes phosphatidylethanolamine, depleting the precursor pool, diacylglycerol, and presumably decreasing the amount of diacylglycerol available for phosphatidylinositol synthesis. The finding that there is an inverse relationship between mouse liver cAMP levels and membrane methyltransferase I activity⁴ (both H-2 influenced) suggests that the MHC affects the two second messenger pathways differently. It is controversial whether coupling of membrane methyltransferases to hormone receptors occurs⁵ but associations of the class I antigen complexes with insulin receptors⁶ and platelet endoperoxide-thromboxane receptors⁷ have been described. The same type of mechanism could be involved in MHC influences on β -adrenergic receptors in man⁸.

ROBERT P. ERICKSON
Department of Human Genetics,
The University of Michigan
Medical School,
Ann Arbor, Michigan 48109-0618, USA

1. Wakelam, M.J.O., Murphy, G.J., Hruby, V.J. & Houslay, M.D. *Nature* **323**, 68–71 (1986).
2. Lafuse, W., Meruelo, D. & Edidin, M. *Immunogenetics* **9**, 57–65 (1979).
3. Lafuse, W. & Edidin, M. *Biochemistry* **19**, 49–54 (1980).
4. Markovac, J. & Erickson, R.P. *Biochem. Pharm.* **34**, 3421–3425 (1985).
5. Hirata, F. & Axelrod, J. *Science* **209**, 1083–1090 (1980).
6. Due, C., Simonsen, M. & Olsson, L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6007–6011 (1986).
7. Curry, R.A., Messner, R.P. & Johnson, G.J. *Science* **224**, 509–511 (1984).
8. Erickson, R.P. *et al.* *Am. J. hum. Genet.* **37**, 124–132 (1985).

Monoclonals and marrow transplants

SIR—E. Donnall Thomas's News and Views article¹ on our paper² may have unintentionally created a misunderstanding about the way in which we suggest monoclonal antibodies could be used to prevent graft-versus-host disease and rejection in bone marrow transplantation. We did not wish to advocate monoclonal antibody treatment nor thymectomy of the donor. Indeed we have been involved in extensive trials using a complement fixing antibody CAMPATH-1 to purge T cells from donor marrow³. In the experimental model we described thymectomy and T-cell depletion of the donor as a convenient way of providing a marrow source purged of T-lymphocytes. May we also add that our own preliminary studies using monoclonal antibody therapy to prevent marrow rejection in man (CAMPATH-1 and CAMPATH-2) have been reported⁴.

HERMAN WALDMANN
S.P. COBBOLD
Department of Pathology,
University of Cambridge,
Addenbrooke's Hospital, Hills Road,
Cambridge CB2 2QQ, UK

1. Thomas, E.D. *Nature* **323**, 110–111 (1986).
2. Cobbold, S.P., Martin, S.Q. & Waldmann, H. *et al.* *Nature* **323**, 164–166 (1986).
3. Hale *et al.* *Bone Marrow Transplant* **1**, 93–94 (1986).
4. Morgan, G. *et al.* *Bone Marrow Transplant* **1**, 107 (1986).

Conquest through biology

Roy Porter

Ecological Imperialism: The Biological Expansion of Europe, 900–1900.

By Alfred W. Crosby.

Cambridge University Press: 1986. Pp. 367. Hbk £27.50, \$22.95; pbk £9.95.

TODAY the developed nations tower over the Third World rather as Gulliver over the Lilliputians. But it was not always so. Indeed the early centuries of European colonial expansion have a David and Goliath quality. How did a few pioneers from a small continent succeed in conquering the oceans, and then all the vast expanses of the Americas, Australasia and much of Africa as well? How did stout Cortés with his handful of men bring the vast Aztec empire of Central America to its knees and wrest its treasures for Spain? How did the flight of a few pious Pilgrim Fathers lead to the decimation of the Red Indians and the opening of a whole continent for the White Man?

The conventional answer to this seeming paradox is that Europeans had overwhelming superiority in science, technology and military might. Caravels, maps, sextants, quadrants, compasses, and above all guns, from the ancient arquebuses to deadly Remingtons, gave the intrepid peoples of the Atlantic seaboard the capacity to come, see and conquer. There is no disputing that, confirms Professor Crosby: could anyone who has ever seen a Western deny that in the long run bows and arrows are no match for rifles, let alone Gatling guns? But the stunning success of European colonization depended on much more than that; in fact, upon a deeper biological and epidemiological "fire power", which Crosby proceeds to explore in this profoundly illuminating and often moving book.

The global success story of the White Man was a true triumph of Darwinian natural selection. When, from the late Middle Ages onwards, Europeans swarmed out from their overcrowded continent, they chiefly put down roots, as might be expected, not at the sterile polar wastes nor in lethal equatorial forests, but in the great open spaces of temperate grassland the globe afforded: almost the whole of North America, much of Middle America, the pampas of South America, the steppes of Siberia, the southern part of

Africa, and Australia and New Zealand. These offered paradise environments, requiring relatively little acclimatization or adjustment of basic European techniques of pastoralism and agriculture. Moreover, these regions of "Neo-Europe" also afforded an ideal biological niche for the invaders.

For, thanks to continental drift, the "splitting of the seams of Pangaea" had left the virgin continents extraordinarily vulnerable. Their natives, such as the Maoris, though often populous, were stuck in the time-warp of the Stone Age and could not compete with European arms, ploughs and liquor. Some — like the Guanches of the Canary Islands (scene of one of the first colonialist triumphs) or the Tasmanians — disappeared altogether from the face of the Earth; others, like the Aborigines, dwindled in numbers to

become shadows of their former selves.

Neither could their flora and fauna compete. The Americas and Australasia did not simply have highly distinct biota. By and large, they specifically lacked powerful and predatory creatures. The pampas with its flightless birds and Australia with its kangaroos were "vacated econiches", wide open and exceptionally hospitable to the horses, pigs, cattle and sheep the White Man brought by design, and to the rats and other vermin he brought by accident. Helped by the introduction of hardy grasses, European livestock, wild and domesticated, thrived in the Neo-Europes, and ensured settlers, fleeing from Malthusian nightmares of starvation in the Old Continent, a luxury seat at Nature's feast.

What ultimately tipped the balance, however, was disease. The colonist and his beasts marched into continents which were amazingly disease-free. But with them they brought all the killer parasites and pathogens associated with filthy, congested, urban Europe, with ports, ships and sailors. The crack troops of the invaders were TB, scrophula, typhus, whooping cough, measles and so forth — microorganisms to which Europeans had developed considerable immunity but which

proved catastrophic to populations exposed to them for the first time. Smallpox destroyed the Amerindians of the pampas, the Aborigines, the indigenes of Siberia; syphilis ruined the Maoris. Before the "magic bullet" came the White Man's "magic bacteria". If nowadays we see our destiny depending on the fight against disease, we do so precisely because, in the past, disease was our ally in assuring the ascendancy of our particular civilization.

Medicine and history, biological and cultural evolution, plagues and peoples are still all too often studied apart from each other, fractured by a stultifying "cultural drift". In his two previous studies, one on the ecological consequences of Columbus's travels, the other on the deadly influenza at the close of the First World War which claimed more victims than the War itself, Professor Crosby has shown that this gap between microbes and men can be bridged. In this latest work, uniting scholarship and speculation, he has unfolded with great power the wider biopolitics of our civilization. □



Facing ruin — a Maori depicted by one of the artists on James Cook's 1768 expedition to the Pacific. (Reproduced from *Ecological Imperialism*.)

Roy Porter is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

Keeping abreast of the revolution

Frank Close

The Forces of Nature, 2nd Edn. By P.C.W. Davies. Cambridge University Press: 1986. Pp.175. Hbk £22.50, \$39.50; pbk £7.95, \$12.95.

The Forces of Nature first appeared in 1979 and established Paul Davies as a popularizer of science. This is probably the best of his non-technical books and the appearance of a second edition is most timely.

When the first edition was completed in mid-1978, particle physics was in the midst of a revolution. The high-energy physicists had recently established that protons, neutrons, pions and other strongly interacting particles are all composites of smaller particles, the quarks. The quantum chromodynamic theory of quark interactions had been formulated and the first quantitative tests of this fundamental theory were being made. The fathers of the theory uniting electromagnetic and weak interactions were about to receive the accolade of a Nobel Prize, even though final confirmation of the theory — the discovery of the W and Z particles — was still four years in the future.

Today the W and Z are known to exist, and the electroweak and quantum chromodynamic theories are referred to as the "standard model" — there being no well-established phenomenon in high-energy physics that is in conflict with this paradigm. For a few months we have opportunity for pause and reflection on what has been achieved in our quest for understanding of the fundamental laws and forces of Nature.

Deeper insights undoubtedly await the experiments at new accelerators that are due to begin at the end of this decade. Physicists are confident that there is far more to be learned because the standard model is not the last word. There are no data in conflict with it, yet there are several theoretical issues on which understanding is hazy. The mystery of the origin of mass is explained in these theories if "Higgs particles" exist; these particles may be produced under conditions obtainable at the new machines. There are many reasons to suspect that the standard model is but a pale glimpse of a more profound unified theory of all forces, including

gravity, which will address questions concerning the origin of matter and the evolution of the Universe from the initial big bang. Indeed, the most significant development since the appearance of Davies's first edition is arguably the realization that high-energy particle accelerators are reproducing in the laboratory conditions similar to those that were prevalent within fractions of a second after the big bang. Thus high-energy experiments have much to tell us about cosmology and the formation of the Universe.

Paul Davies's own contributions have been primarily in general relativity and cosmology. The new symbiosis between high-energy particle physics and cosmology may eventually be seen as one of the great advances in human culture and is discussed in the later pages of this book, though (unfortunately in my view) in relatively little detail compared to the structure of matter and nuclear physics which are covered in the earlier chapters. On the particle physics side of this marriage, Davies describes the phenomenon of CP violation and mentions its importance for

the asymmetry between matter and antimatter in the Universe. He also mentions the search for three families of quarks. Yet I found no mention that CP violation is natural in such a "three generation" universe. This is a surprising omission from the author of *The Accidental Universe*: is this the reason why Nature "requires" three generations when the standard model and life on Earth would, at first sight, be quite happy with just one?

These are minor quibbles when set against much that is good. The most important thing about this book is that you do not need to be a scientist, let alone a physicist, to get a lot from it. Together with its bibliography, it is an important contribution to the growing literature by professional physicists, writing coherently and without recourse to jargon, who wish to share their excitement with a general audience. □

Frank Close is Senior Principal Scientific Officer at the Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK, and Visiting Professor of Physics at Queen Mary College, University of London.

Who's in favour?

M.S. Bretscher

Protein Secretion: A Critical Analysis of the Vesicle Model. By Stephen S. Rothman. Wiley:1985. Pp. 347. £81.25, \$85.

How is it that certain theories or models become accepted, whilst others are rejected? We like to believe that different models consistent with the evidence are proposed, critical tests to distinguish alternatives are performed and eventually one model emerges as closest to the truth, the evidence having been carefully weighed by the scientific community. It is the force of evidence and argument, rather than the personalities of the advocates, that determines the outcome.

In practice, the scientific community does not always weigh the evidence carefully. Today, there is a tendency greater than ever before for the worth of a model to be decided democratically: the more supporters, the better the model. In this democratic process, votes are cast for all to see — the winners are quoted and invited to advertised meetings, and their papers are accepted more readily in the best journals. The leaders of fashion get the laurels and, if there is one person who manages to be seen as the founder of the successful paradigm, he may be elected pope of the field.

Reflecting society at large, science is becoming dominated by media men, who replace reason with glib slogans or polychrome slides. As entertainers fill the

minds of society with a vacuum of trivial news, so the scientific media men are taking over; they can do so because there are too many papers — so many that most never get read at all. Of those papers that receive more than a cursory glance, most only get their titles, abstracts and references scanned. The references, of course, are a poll: they may not tell us which papers are relevant, but rather who is in favour. The plethora of papers arises because of the principle that all observations are novel and must be published: indeed, everyone has a democratic right to publish. The result is that the average paper is really of no interest to anyone and often of little interest to its authors. In this buyers' market, you have to sell, so that everything is hyped. Indeed, good stories are actively fragmented by the authors, the shrapnel being tossed into several journals to create the largest target area, and achieve the greatest chance of a hit.

The reigning paradigm may have much weak evidence in its favour — evidence which in the absence of the model would be regarded as inconsequential. And even if the evidence is against the model — well, its blind adherents will tell you that the model has to be right and the evidence discounted. Behind it all, there grows an infallibility of the pope and his cardinals. This is no new form of mass hysteria, as Peter Mitchell would surely agree.

With this background, Stephen Rothman might identify himself as a heretic, leading a crusade against one of the central dogmas of cell biology: that secreted proteins are forever separated from cytoplasmic proteins by their insertion into the

• Just published in Britain is *Origins: A Sceptic's Guide to the Creation of Life on Earth*, by Robert Shapiro. The book was first published by Summit, New York, and was reviewed by James Lovelock in *Nature* 320, 646 (1986). British publisher is William Heinemann, price is £12.95.

lumen of the endoplasmic reticulum during synthesis. His targets are the pope — in this case, George Palade — and his supporters. The aim of this book is not to defrock anyone; rather, it is to make cell biologists realize that, although vesicles and membrane compartments exist and contain particular components, they are not the transient balloons that people imagine.

Rothman concludes from his own work that the secretory granules found in pancreas are most likely permanent bodies, through whose membranes the contents can diffuse in a regulated fashion. Thus, he argues, the granule contents, including trypsinogen, chymotrypsinogen and amylase, are released from the pancreatic cell by diffusing across the granule membrane into the cytoplasm and, from there, diffusing across the plasma membrane into the extracellular spaces. In other words, the regulated release of proteins from cells is effected by modifying the permeability of these membranes to the secreted proteins, but only to those proteins. Membrane fusion, whereby extracytoplasmic compartments may be joined together, he would like to suggest, is a cell biologist's fantasy. So far, this review must have sounded like music to Rothman's ears, but this is where the music ends.

The first half of the book is concerned with expounding the vesicle model and poking holes in the evidence for it, softening up the reader for the second round. Here, experiments from Rothman's laboratory, carried out over the past 20 years, are described and interpreted: interpreted within the frameworks of his model, which is "simple, physical diffusion", and of the vesicle paradigm involving fusion, which always requires extra *ad hoc* addenda to explain his own experiments and which, to his way of thinking, is unnecessarily complicated. I am afraid it isn't worth discussing the arguments further — an interested reader may wish to look at the evidence. Suffice it to say that Rothman's whole approach is non-molecular; his own experiments are, as he admits in a footnote, sometimes unreliable (p.191); he ignores facts, such as the existence of a glycosylation pathway for secreted proteins and that cytoplasmic proteins are never glycosylated. Moreover, he does not even mention the whole field of endocytosis, which, better than any other, has clearly demonstrated that membrane fusion must occur, and that membrane which goes in does usually come out during exocytosis.

When I go into the booth to vote, I'll spoil my ballot. I care neither for this book, nor for the unquestioning assumption that the majority is always right. □

M.S. Bretscher is Joint Head of the Cell Biology Division in the Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Cellular renewal

David M. Prescott

Molecular Cytology. Vol 1 The Cell Cycle. Vol 2 Cell Interactions. By Jean Brachet. Academic: 1986. Vol. 1 pp. 424, hbk \$59.50, £52; pbk \$34.95, £31. Vol. 2 pp. 512, hbk \$73, £60; pbk \$42, £35.

PROFESSOR Jean Brachet's earlier book, *Biochemical Cytology*, was published in 1957 and it is more than likely that younger researchers in cell biology have never come across it. In its time it was an exciting book that formalized the new, developing way of thinking about cells; cytology fused with biochemistry could produce new insight and understanding about cell structure and function. Contemporary biologists can hardly imagine anything other than such an integrated approach to the cell, but that was not always the case.

Between 1835 and 1957, less than 100 research papers were published on what is now called the cell cycle of prokaryotes and eukaryotes (excluding the mechanism of mitosis). In 1985, over 100 research papers on the cell cycle were published every week. In 1957, a scientist could remain well informed about essentially all phases of what is now called cell biology with a few hours a week in the library; indeed, the term cell biology had then not yet been invented and the *Journal of Cell Biology* was still called the *Journal of Biophysical and Biochemical Cytology*.

The two volumes of *Molecular Cytology* are a second edition of *Biochemical Cytology*, published now 28 years after the first edition. As in the first edition, the author's aim is "to present a critical and synthetic view" of the biochemical basis of cell structure and function. The view is unmistakably that of a biologist who sees biochemistry as a resource for understanding what cells do. Volume I has the subtitle *The Cell Cycle* but only the last 125 pages deal with this subject. The remaining pages are devoted to techniques in cell biology, the cytoskeleton, the plasma membrane, cytoplasmic organelles, chromatin organization, the nucleolus and a few other things. Volume II is subtitled *Cell Interactions*. That means nucleocytoplasmic interactions in unicellular organisms, in somatic cells and in oocytes. But the middle part of the volume is about cleavage, embryonic determination and early development, and the last part discusses differentiation, cell transformation (cancer) and cell ageing. In both books the writing is incisive, and the information is up to date and well referenced.

The author says in the preface that the volumes are intended for advanced students and researchers; that seems about right. The level and style of writing is somewhere between that found in the

Annual Reviews series and the textbook *Molecular Biology of the Cell* by Alberts *et al.* The discussions are enhanced by the author's historical perspective, some of it the fruit of personal experience. History in science doesn't appeal to everyone, especially not to the young. It's a taste acquired as one becomes a part of history.

The coverage of material in these two volumes is a remarkable accomplishment for a single individual; it represents many years' accumulation of knowledge and provides the reader with an integrated view that multi-authored books never achieve. No book is ever perfect, and these volumes have their imperfections. Much more is known about the causes of cell transformation (cancer) than the author states; amitosis is not "an agonic phenomenon" but alive and well in ciliated protozoa; and G1 of the cell cycle is not generally assumed to be a period of accumulation of machinery required for DNA synthesis.

Overall, however, the two volumes achieve their purpose. There is little in them that I could disagree with and they are worth the attention of students and researchers. I expect to refer to them regularly for some years to come. □

David M. Prescott is Distinguished Professor in the Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, USA.

THE MANAGEMENT OF AIDS PATIENTS

Edited by DAVID MILLER, JONATHAN WEBER and JOHN GREEN

The Management of AIDS Patients is the first comprehensive guide to the practical clinical management of patients with AIDS or HTLV III infections. The book avoids the sensational aspects of the disease, offering solid advice and information for all people involved in patient care.

The editors and many of the contributors come from St Mary's Hospital, London one of the foremost centres in Britain for the treatment of AIDS patients. The knowledge and experience of these experts has been combined to provide a much-needed book for doctors, nurses, dentists and indeed for health-care professionals everywhere.

1986
200pp Hardback 0 333 40465 3 £30.00
Paperback 0 333 40466 1 £10.95

Please order this title from your bookseller — or in case of difficulty from Judy Chappell, Macmillan Press, Houndmills, Basingstoke, Hants RG21 2XS.

M MACMILLAN PRESS

"The most
important science
book of the year"

Good Book Guide

THE BLIND WATCH- MAKER

by
Richard Dawkins
author of *The Selfish Gene*

"It succeeds quite brilliantly
... again and again, it brings
home the nature and force of the
central evolutionary mechanism
of natural selection in a way that
I have never seen or felt
previously"

Michael Ruse

£12.95

Cased Illustrated
ISBN 0 582 44694 5

Longman
Scientific &
Technical

Fractal attractions

Michael Berry

The Beauty of Fractals: Images of Complex Dynamical Systems. By H.-O. Peitgen and P.H. Richter. *Springer-Verlag*:1986. Pp.199. DM 78, \$29.50.

THE surprises by which discovery proceeds in science and mathematics usually turn out to have been foreseen by somebody, albeit dimly. But until very recently not even a science-fiction writer would have anticipated this picture book celebrating the incredibly rich geometry encoded in the humble quadratic equation.

The complexity arises from a mathematical procedure so simple that it can be explained in words. Take a point in the plane, and represent it by a complex number. Square this complex number and add to it a (complex) constant, thereby generating a new point. Repeat the process over and over again. Depending on the starting point, the iterated points may escape to infinity, or they may not. The boundary between these two sorts of behaviour is called the Julia set. What is surprising is how intricate the Julia sets can be, and how delicately they depend on the value of the constant that is added at each step. As Adrien Douady puts it:

Some are a fatty cloud, others are a skinny bush of brambles, some look like the sparks which float in the air after a firework has gone off. One has the shape of a rabbit, lots of them have sea-horse tails. . . .

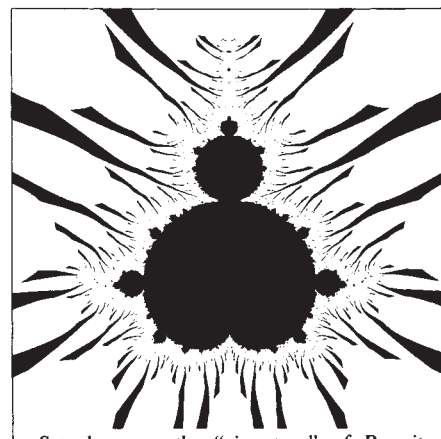
Julia sets are fractals, that is geometric objects with structure down to infinitely fine scales, indicated by self-similarity under magnification. For some values of the added constant, the Julia set is a connected fractal, that is an infinitely crinkly curve; for others, it is a fractal dust, that is a cloud of disconnected points. The boundary between these two sorts of behaviour defines another amazingly complicated set; this is the Mandelbrot set, which lives in the plane whose points label the (complex) values of the added constant.

Peitgen and Richter present high-resolution computer-generated images of the Julia and Mandelbrot sets, vividly coloured to illustrate their important mathematical properties (for example the contours of electrostatic potential if the sets are charged conductors). The pictures are startlingly beautiful, especially the dizzying sequence of panoramas zooming into the Mandelbrot set until part of it is magnified a million times. Originally the pictures formed a travelling exhibition. Here they are accompanied by an explanatory text and several invited essays.

Introducing the explanations, the authors emphasize the connection between

fractals and chaos: on the Julia set, for example, the moving point (which just fails to escape) wanders irregularly forever. They argue with eloquence and clarity that our scientific view of the world is altered by the realization that a simple procedure can generate both infinitely complex patterns and randomness. It helps dissolve an apparent contradiction: the fundamental equations of physics, which are supposed to describe the world accurately, are simple and deterministic, but the world is manifestly complicated and unpredictable.

The mathematical explanations are not so clear. My suspicion is that the authors,



Set shape — the "signature" of Benoit Mandelbrot.

being physicists, have strained too hard to ensure that their formulations are mathematically impeccable. By contrast, the informal exposition by Douady (who is a mathematician) is a masterpiece.

If the Julia and Mandelbrot sets arose only in quadratic equations they would be interesting merely as mathematical freaks. In fact they are encountered again and again in non-quadratic contexts, and this ubiquity ("structural stability") is what makes the discovery of their geometry so important. Gert Eilenberger's thoughtful essay takes up this point and addresses the

... emotional question: why is it that the products of our technology, ... seem to be so unnatural when they are products of natural science?

His answer is that

our technical products are made stiff by the complete orderliness of their forms and functions ... [which] is *untypical* even for quite 'simple' natural processes.

Fractal geometry in its modern form is the single-handed creation of Benoit Mandelbrot. Here he contributes useful reminiscences, mainly about the set that bears his name.

A wide variety of specialist and non-specialist readers will derive pleasure and instruction from this gorgeous book. □

Michael Berry is a Professor in the H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.

Is Britain spending enough on science?

from John Irvine and Ben R. Martin

Britain is falling behind the major industrialized countries in its investment in academic and related research, which may explain the nation's declining contribution to world science.

In a previous article¹, we presented a range of publication and citation-based indicators of research output which lent support to the view that British science is in a state of decline, if not crisis. This view is rapidly becoming conventional wisdom, and will be further strengthened by a Royal Society report published today on the health of basic research in Britain, a summary of which will appear in *Nature* next week². Before the problem can be rectified, however, the factors underlying the decline need to be understood. Several possible explanations might be advanced: are we in Britain obtaining poor value for money from existing investments as a result of inefficiencies in research-performing organizations? Are we perhaps continuing to channel funds into mature fields with little potential for major scientific advance? Or is the overall level of funding inadequate relative to other industrial nations? To provide better background information for considering these questions, the Advisory Board for the Research Councils (ABRC) has in recent years initiated a programme of work on science policy. Here we summarize the findings of a study, commissioned from the Science Policy Research Unit as part of that programme, on 'An International Comparison of Government Funding of Academic and Academically Related Research'³.

Objectives

The study's overall objective was to compare British government expenditure on academic and related research in 1975, 1980 and 1982, with that of France, Japan, the Netherlands, the United States and West Germany. As well as overall totals, expenditure data were compiled for 9 main fields (for example, physical sciences) and 40 subfields (such as chemistry, nuclear physics and astronomy). Figures at this level of disaggregation were required to address such policy-related questions as the following: do patterns of expenditure by field and subfield differ markedly across countries? Do some nations invest relatively more in academically related establishments (such as Max-Planck institutes or research council laboratories) as opposed to university-based research? What trends have occurred in the level and distribution of expenditure, and the types of research support mechanisms employed?

Before outlining the results, it is first

necessary to comment on the limitations of existing statistics on research and development (R & D) expenditure. One factor behind ABRC's decision to fund the collection of new figures was the difficulty in using the official statistical series published by the Organization for Economic Cooperation and Development (OECD)⁴, the main source of financial data on higher education research. Four main sets of problems can be identified with the data reported to OECD by member states:

(1) Coverage of the 'higher education' sector varies widely across countries. Only expenditure in universities and university-level 'technical high schools' is included for the Netherlands, for example, while for Japan the sector comprises more than 1,050 organizations, over half of which are two-year junior and technical colleges. Similarly, research council laboratories are excluded for the Netherlands and Britain (being classified as part of the 'government' sector), while the equivalent CNRS establishments are included for France, and the US 'higher education' sector even extends to large mission-oriented centres such as Los Alamos. Further inconsistencies arise from uneven coverage of the health-care component of university hospitals, and from the exclusion of all expenditure on international facilities like CERN even though equivalent national laboratories are included.

(2) Technical problems exist in the methods used to estimate expenditure on infrastructural support for academic research. Since such support normally comes from the general funds provided to universities for both teaching and research, coefficients derived from surveys of how academic staff divide their time between teaching and research are generally employed to estimate the research component of those funds. The latter, for which OECD uses the term 'general university funds' (GUF), constitutes by far the largest element of academic research expenditure in most countries — as well as being the chief source of error in the overall figures. Errors stem principally from (i) the use of obsolete time-budget data (the British figures, for example, are 16 years old), (ii) poorly constructed research coefficients (the figure of 30% used for British universities excludes various re-

search-related activities classified under 'unallocable internal time' which, if included, have the effect of raising the coefficient to 42%), and (iii) in at least one case from political intervention (a small increase in the research coefficient can result in quite a large apparent rise in government expenditure on university research).

(3) The level of disaggregation of the data is insufficient for most policy purposes — OECD publishes figures for only five main fields (one of which, for example, covers the whole of natural sciences), with just three being reported for some countries.

(4) International comparisons of different types of funding mechanism are difficult because of the inconsistent treatment of GUF and 'direct government funding' of higher education research. This a particular problem for France and Japan.

The above sources of error are far from insignificant with the result that many of the data are not fully internationally comparable — the OECD official figure for Japanese GUF in 1982, for example, is 50% higher than that estimated in our study after applying common criteria for coverage across the six countries. As a result, 'bids' by the scientific community for increased funding made on the basis of comparison with international norms can all too easily be dismissed by those disinclined to accept them simply by pointing to the undoubted inadequacies of the OECD statistics. Informed debate on the funding of higher education research consequently requires better expenditure data. While OECD officials are fully aware of this and are giving priority to the problem (they provided much assistance in our study), their efforts have to date met with little response from member states.

Measuring expenditure

The growing strategic importance of fundamental research is resulting in increased emphasis being placed everywhere on improved management and planning within the academic sector. It was, therefore, relatively easy to secure the participation of relevant authorities in each country for a study aiming to provide better statistical information. The first step was to obtain

preliminary agreement on coverage and the best way of compiling the necessary figures. For each country, our aim was to obtain disaggregated data relating to three expenditure categories: (a) academic research financed by GUF; (b) academic, separately budgeted research (ASBR), taking the form of peer-reviewed grants and other funding earmarked for research; and (c) academically related research (ARR) undertaken outside universities in (i) national and international laboratories providing central facilities for use primarily by academic researchers, and (ii) institutes financed mainly through government core-funding (as opposed to commissioned research) and engaged in longer-term and more basic research similar to that found in academic institutions in most other countries (these are generally research council establishments but also include various agricultural and medical laboratories). To some extent, ARR can be seen as a residual category enabling us to bring into the study functionally equivalent organizations formally outside the higher-education sector and hence ensure that 'like' was being compared with 'like'. ARR coverage was revised progressively in line with the comments of officials, so that by the end there was, with few exceptions, general agreement on coverage.

Early effort also focused on constructing an internationally acceptable field and subfield breakdown which accommodated the different research traditions and institutional structures of each country. (This involved compiling a comprehensive thesaurus of specialities coming under each subfield.) After consultation and piloting, a modified and extended version of the US National Science Foundation research classification scheme was adopted.

Data collection

The next stage was to determine appropriate methods of data collection. For GUF, the normal procedure was to rework published expenditure figures in collaboration with relevant national authorities. This often entailed using subsidiary unpublished data prepared specially for the project (for example, the organizations which assisted in Japan were Monbusho and the Statistics Bureau). In the case of ASBR, for the four European countries we sent our questionnaires to agencies providing research grants to higher education establishments. Completing the survey forms generally involved these organizations in substantial work reclassifying expenditure categories in line with those used in the study and redistributing administrative costs proportionately. The most difficult task, however, was to obtain figures for ARR since with the exception of the UK this involved contacting large numbers of disparate agencies and laboratories, and in certain cases making somewhat arbitrary

		1975	1980	1982	% change 1975-82*
General university funds (GUF) only	Neths	550	600	620	13
	UK	890	920	990	11
	France	480	680	840	75
	FRG	2,050	1,930	1,930	-6
	Japan	1,500	2,090	2,240	50
	US	1,510	1,770	1,880	25
Total academic (GUF + ASBR)	Neths	640	730	740	17
	UK	1,180	1,290	1,380	16
	France	1,050	1,310	1,620	55
	FRG	2,580	2,470	2,460	-5
	Japan	1,940	2,810	2,980	55
	US	7,380	8,180	8,220	11
Total academic and academically-related research	Neths	780	900	900	15
	UK	1,820	1,830	1,930	6
	France	1,740	2,120	2,590	50
	FRG	3,340	3,290	3,300	-1
	Japan	2,070	3,010	3,170	55
	US	8,310	9,370	9,370	13

* After rounding

		1975	1980	1982	% change 1975-82*
General university funds (GUF) only	Neths	40.0	42.5	43.5	9
	UK	16.0	16.5	18.0	11
	France	9.0	12.5	15.5	70
	FRG	33.0	31.5	31.5	-5
	Japan	13.5	18.0	19.0	40
	US	7.0	8.0	8.0	14
Total academic (GUF + ASBR)	Neths	46.5	51.5	52.0	12
	UK	21.0	23.0	24.5	17
	France	20.0	24.5	30.0	50
	FRG	41.5	40.0	40.0	-4
	Japan	17.5	24.0	25.0	45
	US	34.5	36.0	35.5	2
Total academic and academically-related research	Neths	57.0	63.5	62.5	10
	UK	32.5	32.5	34.5	6
	France	33.0	39.5	47.5	45
	FRG	54.0	53.5	53.5	-1
	Japan	18.5	26.0	26.5	45
	US	39.0	41.0	40.5	4

* After rounding

judgements about what expenditure to include.

Finally, the data for each country were processed and draft tables with an accompanying commentary sent to participating bodies for comment. In nearly all instances, proposed revisions were fully accommodated. In order to permit comparison of spending in real terms over time, data for all three years were expressed in terms of 1982 costs using the OECD deflator for GDP. It should be stressed that such a deflator may significantly understate the rise in costs for R&D, the effect being largest for countries like the United Kingdom with an above average rate of inflation. Because exchange rates fluctuate and tend not to represent actual purchasing power within the different

countries, we employed the 'purchasing power parity' exchange rates developed at OECD to convert national currencies to 1982 US dollars⁵.

Results

Let us now turn to the results, examining first the figures on total government-funding of academic and academically related research, and the main trends between 1975 and 1982. We then focus on the expenditure breakdown by main field in 1982. Finally, data are presented for selected subfields, again in 1982.

Table 1 shows the distribution of government expenditure in each country between GUF, ASBR and ARR. In the United Kingdom GUF grew by 11% in real terms during the seven years up to 1982

Table 3 Government funding as % of GDP

		1975	1980	1982	% change 1975-82*
General university funds (GUF) only	Neths	0.430	0.415	0.440	2
	UK	0.190	0.180	0.195	2
	France	0.100	0.120	0.145	45
	FRG	0.355	0.285	0.290	-19
	Japan	0.165	0.180	0.180	8
	US	0.059	0.058	0.061	5
Total academic (GUF + ASBR)	Neths	0.500	0.505	0.525	5
	UK	0.255	0.255	0.270	6
	France	0.215	0.230	0.280	30
	FRG	0.450	0.365	0.365	-18
	Japan	0.215	0.240	0.240	12
	US	0.290	0.270	0.270	-6
Total academic and academically- related research	Neths	0.615	0.620	0.635	3
	UK	0.390	0.365	0.375	-3
	France	0.355	0.370	0.445	25
	FRG	0.580	0.485	0.490	-15
	Japan	0.230	0.260	0.250	11
	US	0.325	0.310	0.310	-5

* After rounding

Table 4 Government funding in 1982 broken down by field (1982 \$M)

	Neths	UK	France	FRG	Japan	US
Engineering	90	300	210	410	680	1,140
Physical sciences	180	420	840	840	450	1,420
Environmental sciences†	20	130	120	150	100	590
Maths and computing	30	90	140	110	70	280
Life sciences	350	660	940	1,190	1,050	4,630
Social sciences	100	110	120	160	130	590
Professional & vocational	60	100	60	180	330	270
Arts and humanities	70	120	170	230	340	290
Other (multi-disciplinary)	-	-	-	20	-	170
Total*	890	1,930	2,590	3,300	3,170	9,370

† These figures are not strictly comparable because of differences in the way some aspects of environmental research, in particular geological surveying, are funded in individual countries.

* Details may not sum to totals because of rounding.

from \$890M (1982 dollars) to \$990M. Even then, it was only approximately half that for Germany, Japan and the US, while GUF in the Netherlands in 1982 was two-thirds the British figure, despite having one quarter the population. French GUF is relatively low, because a greater proportion of support for academic research is separately budgeted than in other European countries. For all academic research (GUF plus ASBR), the French figure in 1982 was nearly 20% higher than that for the UK (\$1,620M), the latter being only just over half the West German and just under half the Japanese totals. However, all these were dwarfed by the US which in 1982 spent almost as much as the other countries together. A similar picture emerges for academic and academically related research combined: UK expenditure was only just over twice that for the Netherlands; the French spent 35% more than Britain, the Japanese and Germans 65-

70% more, while the United States total was nearly five times greater.

Differences

Clearly, the above statistics fail to take account of differences in size and wealth between the countries. The two methods most commonly used to correct for this are to represent research expenditure in per capita terms and as a percentage of gross domestic product (GDP). Results for the former are shown in Table 2. For academic research (GUF plus ASBR), the highest spenders in 1982 were the Dutch with \$52 per head of population, followed by Germany and the US, with Britain the lowest of the six countries. When academically related research is taken into account, the UK spent less than all the other countries apart from Japan. (Japan does have a considerably larger civil science budget, but channels a higher percentage into R&D carried out in laboratories outside the academically related sector,

for example those coming under the Science and Technology Agency.)

Table 3 shows the expenditure totals expressed as a percentage of GDP. For all academic and academically related research combined, the United Kingdom, with a figure of 0.375% of GDP in 1982, was some way behind France (0.445%), Germany (0.490%) and the Netherlands (0.635%), although it was ahead of the United States and Japan, the two countries with the largest GDPs. In addition, whereas the UK figure fell by 3% between 1975 and 1982, that for France grew by 25%.

In short, however the results are presented, it is difficult to avoid the conclusion that the United Kingdom spent less overall on academic and related research than its two closest European competitors, France and Germany. Furthermore, the gap which opened up between France and Britain over the period studied is now almost certainly wider since government support for research grew much faster in France during 1982-86.

Distribution

Let us next consider how the expenditure for each country was distributed among the main fields. Table 4 contains data on academic and academically related research for 1982. In the case of engineering, the UK spent \$300M, just over 15% of the national total (\$1,930M). This compares with \$410M for Germany, \$680M for Japan, and over \$1,100 million for the United States. It is significant that Japan devotes a much larger proportion (over 21%) of its overall effort to this field than the other countries. Conversely, the small French figure reflects the relatively low priority given to engineering until recently.

Turning next to physical sciences, Table 4 shows that Germany and France spent twice as much as Britain in 1982, and the United States over three and a half times more. Japan, however, was only slightly ahead of the United Kingdom, physical sciences accounting for just 14% of the Japanese total compared with an average of over 20% for the other five countries.

For environmental sciences, the comparatively high UK figure — intermediate between that for Germany and France — partly reflects the fact that research council funds are used to support activities like the Geological Survey which in other countries are not part of the academically related sector. For mathematics and computing, Britain was a little behind Germany, this being a field especially well supported in France. Britain was somewhat less generous towards the life sciences, with the French spending over 40% more and the Germans 80% more. However, the gap between Britain and its two main European competitors was smaller in social sciences, an area given

Table 5 Government funding of selected subfields in 1982 (1982 \$M)

	Neths	UK	France	FRG	Japan	US
Chemical engineering	18	20	17	14	110	100
Electrical engineering	19	65	65	80	150	250
Mechanical engineering	18	45	15	75	120	150
Chemistry	40	120	200	230	85	390
Physics	100	200	550	500	260	800
Computer sciences	12	35	50	45	—*	150
Biological sciences	60	250	650	330	140	1,700
Agricultural & veterinary	45	65	40	150	190	700
Medicine/health	230	340	230	700	700	2,000

* data not available

considerable emphasis by the Dutch.

Table 5 shows government expenditures in 1982 on various subfields. For chemical engineering, UK spending was relatively large compared with France and Germany, as was that for the Netherlands given the country's size. (The high Japanese figure may stem from definitional differences over the distinction between 'chemical engineering' and 'chemistry' — as can be seen from the table, funding for chemistry was rather low.) For electrical and electronic engineering, Britain and France were a little behind Germany, all three investing much less than Japan and the United States. The latter two also spent by far the largest amount on mechanical engineering, a subfield where the British total was only 60% of that for Germany. (The small French figure again possibly reflects definitional differences as much as low priority for the subfield.)

The gap in levels of expenditure between Britain and its two main European rivals is particularly large for chemistry and physics. For the former, the French devoted 65% more in 1982 and the Germans over 90% as much. In physics, France and Germany each apparently spent no less than two and a half times as much as Britain, the gap having widened significantly since 1975 when French support was only 50% greater than in the United Kingdom. Even in computer sciences, an area identified by policy-makers as deserving priority for strategic reasons, British expenditure lagged appreciably behind that of France and Germany.

The data for the three life sciences shown at the bottom of Table 5 are more difficult to interpret because of possible definitional differences, with some expenditure classified in France under 'biological sciences' consisting of work that in other countries would be regarded as agricultural or health research. However, it is significant that in all three life sciences, Britain spent much less than Germany which was, in turn, a factor of four to five behind the United States.

Overall findings

Before drawing conclusions, a word of caution is needed on the reliability of the expenditure data reported, some of which (in particular those for agricultural, health and environmental research) must be treated with great care. In order to achieve the requisite degree of disaggregation into the 9 fields and 40 subfields in our research classification scheme, we sometimes had to resort to that stock-in-trade of the statistician — 'pro-rating' (breaking down an item of expenditure into its component parts pro rata to some other variable for which suitably disaggregated data do exist). In several cases, this involved making a whole series of assumptions which, although the 'best' possible within the boundary conditions for the study, were little more than crude approximations. This said, it is encouraging that officials in the six countries who have read the report see our estimates as a significant improvement on earlier international comparisons of academic and related research funding. In particular, the study seems to have achieved greater international consistency as regards the coverage of the academic (and to a lesser extent the academically related) research sector. It has also succeeded to some degree in disentangling the main mechanisms of government support for such research so that one can now see the relative emphasis given by different countries to infrastructural support financed by GUF compared with separately budgeted research. Finally, it has yielded considerably more disaggregated statistics.

Conclusions

What conclusions, then, can be drawn from the results? From the point of view of the United Kingdom, it is clear that overall government funding of academic and related research is falling increasingly behind that of our nearest European competitors — especially if account is taken of our use of a non-R&D deflator. Furthermore, the gap is especially large in cer-

tain fields with considerable technological potential such as physics, computer sciences and biology.

How concerned should British policy-makers be by this? It could quite justifiably be argued that national research performance also depends on factors other than the level of funds available. The relationship between the inputs to, and outputs from, science is one where further work is much needed. Up to now, such work has been hindered by the absence of satisfactory (in particular, internationally comparable) data on both research expenditure and scientific outputs. However, on the basis of some preliminary analysis a year ago, we pointed to the likely existence of a link between Britain's spending on academic and related research and the country's changing standing in world science¹. Although it is not the purpose of the study reported here to analyse the relationship between scientific inputs and outputs, it is surely significant that physics — one of the areas where the gap in expenditure between the UK and its principal rivals grew most rapidly between 1975–82 — is also one where Britain's world-share of publications has fallen most sharply (declining by 21% over the period 1973–82). Other factors notwithstanding, it can scarcely be denied that the ability to attract and retain the brightest minds and to purchase state-of-the-art research equipment must depend to a large extent on the availability of funds. If so, Britain's hopes of remaining a leading scientific power can only be described as slim in the light of the results from this international comparison of funding.

We gratefully acknowledge financial support from the Advisory Board for the Research Councils, the Leverhulme Trust, and the Economic and Social Research Council, as well as statistical assistance provided by Nigel Minchin. □

John Irvine is with the Technical Change Centre, 114 Cromwell Road., London SW7 4ES, and Ben Martin is a Fellow of the Science Policy Research Unit, Falmer, Brighton BN1 9RF.

1. Irvine, J., Martin, B., Peacock, T. & Turner, R. *Nature* **316**, 587–590 (1985).
2. Smith, D.C., Collins, P.M.D., Hicks, D.M. & Wyatt, S. *Nature* **323**, 681–684 (1986).
3. Martin, B.R., Irvine, J. & Minchin, N., *An International Comparison of Government Funding of Academic and Academically Related Research — Vol.1 (Main Report) and Vol.2 (Appendices)*, ABRC Science Policy Studies No.2 (Advisory Board for the Research Councils, London, 1986 — available only from SPRU).
4. *OECD Science and Technology Indicators: Resources Devoted to R&D* (Organization for Economic Co-operation and Development, Paris, 1984).
5. Ward, M. *Purchasing Power Parities and Real Expenditures in the OECD* (Organisation for Economic Cooperation and Development, Paris, 1985).

The superstring: theory of everything, or of nothing?

John Ellis

Theory Division, CERN, CH-1211 Geneva, Switzerland

Superstring models excite theoretical physicists because they may unite the four fundamental forces. These theories are formulated in a ten-dimensional world of extraordinarily high energies. Recent work indicates how superstrings may nevertheless relate to our four-dimensional world and to laboratory experiments.

THE corridors of theoretical physics institutes are buzzing these days with talk of cohomology. There is a run on the mathematics texts in the physics libraries. The cause of all the excitement is the superstring¹, which many particle theorists are now taking seriously as a candidate 'theory of everything' (TOE) which may unify all the fundamental interactions—electromagnetism, strong and weak nuclear forces and gravity—and explain the number and couplings of all the elementary particles. String theories have existed for about 15 years². According to the superstring theory, all the particles hitherto considered as 'elementary' are in fact modes of oscillation of a one-dimensional extended object in ten space-time dimensions. Here I try to explain why a long-observed, radical and apparently bizarre theory has recently caught the imagination of many particle theorists. There is now a favoured strategy for extracting four-dimensional physics at accessible energies, although it has some doubtful aspects. Some prospects for experimental tests of the superstring exist, and some critiques of superstring mania have appeared³. Although there is still a yawning gap between the hopes for the superstring and its achievements, its audacious claim to explain all of fundamental physics cannot be ignored.

Let us start by reviewing the context into which the superstring has erupted. Since the discovery of the weak gauge bosons $W^\pm$ and Z^0 at the CERN proton-antiproton collider in 1983, no particle physicist seriously questions the basic validity of the Standard Model³, which comprises quantum chromodynamics (QCD) for the strong interactions and the Glashow-Weinberg-Salam model for the electroweak interactions. There is equal unanimity that the Standard Model is inadequate and must be extended to resolve several basic problems. These can be grouped into three general areas, those of unification, flavour and mass. The unification problem is that although the fundamental particle interactions have similar structures based on gauge theories, they are described by distinct $SU(3)$, $SU(2)$ and $U(1)$ gauge groups with independent gauge couplings—hardly Einstein's vision of a unified field theory. The flavour problem is that of understanding the apparent proliferation of species (flavours) of fundamental matter particles—at least six quarks and six leptons—as well as the apparently random ratios of their masses and the relative strengths of their various charged-current weak interactions, normally described by the Cabibbo-Kobayashi-Maskawa mixing angles. The mass problem is that of understanding the origins of all these quark and lepton masses, as well as those of the $W^\pm$ and Z^0 , and of understanding why they are so much smaller than the other apparently fundamental mass scale in physics, the Planck mass $m_P \approx 10^{19}$ GeV associated with gravitation. (The universal gravitational constant G is equal to m_P^{-2} .)

A favoured approach to the unification problem has been that of grand unification: the disparate strong, weak and electromagnetic interactions are combined into a simple gauge group with a single gauge coupling. In order to accommodate the great disparity between the strong and electroweak couplings, this grand unification can take place only at dizzying energies of $\sim 10^{15}$ GeV. Grand unified theories (GUTs) have had some

circumstantial successes, notably in calculating the neutral weak current mixing angle θ_w and in predicting the mass of the bottom quark. However, none of the dramatic new phenomena advocated by GUTs, such as proton decay, magnetic monopoles or neutrino masses, has ever been convincingly observed. Although these failures cast doubt on some schemes of grand unification, the general idea cannot be disproved, as the long extrapolations to high energies and theoretical uncertainties do not permit reliable predictions to be made.

One approach to the flavour problem is to postulate that the apparently elementary quarks and leptons are actually composite objects containing more fundamental particles called preons, bound together by new superstrong interactions. There is as yet no experimental evidence for such a substructure—no quark form factors or excited leptons—which implies that the intrinsic sizes of quarks and leptons must be smaller than 10^{-17} cm, associated by the uncertainty principle with energies above 1 TeV. It is very difficult to make a consistent theory whose bound states have masses many orders of magnitude smaller than their binding energy, and many such models of preon dynamics are as complicated as the quarks and leptons which they claim to explain. Such composite models therefore have little attraction for me, especially as they strike me as philosophically very naive—yet another layer of the cosmic onion.

The mass problem is tackled in the Standard Model by postulating an elementary spin-zero Higgs boson H , whose couplings to other particles g_{Hf} determine their masses: $m_f \propto g_{Hf}$. The theoretical consistency of the Weinberg-Salam model requires the existence of at least one physical Higgs boson with mass ≤ 1 TeV, and a primary task of the LEP accelerator now under construction at CERN will be to produce and observe it if it weighs less than ~ 100 GeV⁴. Unfortunately, it is very difficult to predict the mass of the Higgs boson: the theory tends to acquire corrections which are much larger than any reasonable physical value, and may be as large as the grand unification scale or the Planck mass. One proposal for taming these corrections was to make the Higgs boson composite⁴, but theories based on this idea had trouble explaining quark and lepton masses and are now disfavoured. An alternative way to stabilize the Higgs mass is to postulate a complete set of supersymmetric partners for the known particles, which cancel out the excessive corrections to the Higgs mass⁵. To do this job, the supersymmetric particles must have masses of ≤ 1 TeV. Searches for them at masses up to 50 GeV have been unsuccessful, despite some false alarms⁵. If supersymmetry is made local; that is, if one allows supersymmetry transformations to be different at different points in space, just like conventional symmetries in gauge theories, one must include gravity and construct a supergravity theory. Many theorists have thought for some time that this is the appropriate framework for discussing particle physics.

The role of dimensions

Simple supergravity by itself is not a unified theory, and has nothing to say about the unification and flavour problems,

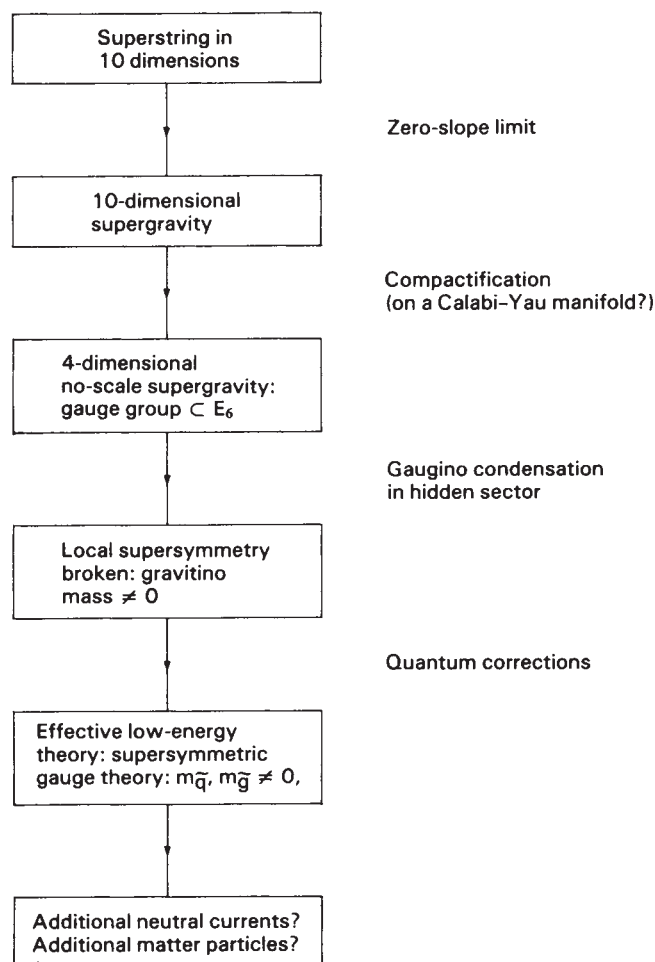


Fig. 1 Sketch of the strategy for making contact between the superstring theory and reality.

although it can be combined with the approaches to these problems which have already been mentioned. To make supergravity itself more unified, one should either postulate more supersymmetries, or reformulate the theory in a higher number of space-time dimensions, or both. Unfortunately, no one knows how to make a theory with more than one supersymmetry discriminate between left and right, that is, to violate parity as observed in the weak interactions. It is also difficult to make a theory originally formulated in more than four dimensions violate parity when it is reduced to the four physical dimensions, but it can be done if one starts from a theory containing gauge fields in an even number of dimensions. The largest even number of dimensions in which a supergravity theory can be formulated is ten, so the ten-dimensional supergravity theory⁶⁻⁸ provides a natural starting point for unifying all matter and forces. Unfortunately, ten-dimensional supergravity seemed to make no sense as an ultimate TOE because its quantum corrections had uncontrollable infinities and violated the gauge symmetry of the theory—in the jargon of the trade, it was unrenormalizable and had anomalies. It was at this stage that Green and Schwarz⁹ came to the rescue with the superstring.

String theories have been around for about 15 years, and were originally conceived as theories of strongly interacting particles². They went into eclipse around 1974 as QCD became accepted as the fundamental gauge theory of the strong interactions and it was realized that bosonic strings (superstrings) could be formulated consistently only in 26 (or 10) space-time dimensions^{10,11}. At about the same time, however, it was realized that the massless particles of a ten-dimensional superstring

theory had the same interactions as the graviton, conventional (Yang-Mills) gauge particles and their supersymmetric partners in ten dimensions^{12,13}, and the suggestion was made¹⁴ that the superstring should be interpreted as a quantum theory of gravity. However, it was by no means obvious that the superstring theory was renormalizable, let alone finite, and it was thought to have anomalies in general. The great contribution of Green and Schwarz⁹ was to discover that there were no infinities or anomalies when one calculated one-loop amplitudes in the superstring with either of two particular internal gauge groups— $O(32)$ or $E_8 \times E_8$. Their discovery gave hope that the superstring could be taken seriously as a theory of quantum gravity and TOE which is finite, although this has yet to be proved.

Particles in the superstring theory are modes of oscillation of a one-dimensional object—a string—in one time and nine space dimensions. This is not a traditional composite model with point-like constituents. The full spectrum contains infinitely many excited states whose squared masses are quantized in units of m_P^2 . It is supersymmetric and the lowest harmonics, which have no mass in a first approximation, have all the quantum numbers and interactions of the ten-dimensional supergravity theory mentioned above. Evidently the formulation of the superstring is several steps away from physics in four dimensions at energies ≤ 100 GeV, which is where finance and technology confine our current experiments.

The role of topology

Many particle theorists have adopted a strategy for making contact with reality which is illustrated in Fig. 1 and relies heavily on the mathematics texts referred to above. Conventionally, one has first discarded the excited modes of the superstring theory, obtaining the ten-dimensional supergravity theory mentioned above, supplemented in principle by an infinite sequence of higher-order interaction terms with coefficients of order m_P^{-n} . For historical reasons this procedure is known as taking the zero-slope limit. The next step after arriving at the 10-dimensional supergravity theory has been to reduce the number of space-time dimensions to four by 'compactifying' the six surplus dimensions. Progress is now being made on compactifying the superstring theory directly to a four-dimensional supergravity theory, thus bypassing the intermediate ten-dimensional supergravity stage. The type of manifold with which one curls up these extra dimensions is constrained by three principal criteria¹⁵. The four-dimensional space-time should be maximally symmetric, so that it has a chance to describe the Universe we see; supersymmetry should be maintained, so that the mass problem may be tackled in the manner described previously; and the effective four-dimensional theory should violate parity.

A favoured solution¹⁵ to the first two of these conditions is to compactify on what mathematicians call a Calabi-Yau manifold^{16,17}. This is a compact space with six real dimensions that can be paired into three complex coordinates. It has the special property that at each point in the space the theory is invariant under unitary rotations in a three-dimensional complex space. These rotations form a group called the holonomy group, which is $SU(3)$ for a Calabi-Yau manifold. The third condition seems to rule out superstring models with an $O(32)$ internal gauge group: only $E_8 \times E_8$ models can yield interactions in four dimensions that violate parity. Starting from this group, Calabi-Yau compactification yields a four-dimensional gauge group which is some subgroup of E_6 , one of the traditional GUT groups, together with matter particles occurring in sets of 27 that are related to each other by E_6 symmetry. Each of these sets of 27 particles contains the quarks and leptons of a conventional GUT generation, supplemented by some additional particles including another quark of charge $-1/3$, two Higgs multiplets, and a pair of neutral fields, of which one has all the properties expected of a right-handed neutrino. We will return later to the phenomenology of these additional particles. The number of such 27-dimensional E_6 generations is determined

by topology, specifically by the number of holes in the manifold, which is the Euler number of the space of compactification¹⁵. There are in general other low-mass fields, whose number is determined by other topological quantities, called Betti numbers, whose physical interpretation is more subtle. Although Calabi-Yau manifolds are often preferred, there may be other phenomenologically interesting manifolds of compactification which could lead to SO(10) or SU(5) GUT groups¹⁰. In these cases also, topological numbers would specify the numbers of light matter species.

Topology also provides a novel mechanism for breaking the gauge symmetry^{15,19-21}, which is a generalization²² to non-Abelian gauge theories of the Bohm-Aharonov effect. If the compactification manifold has suitable 'holes', the gauge fields on the manifold may make it look as if there were a tube of non-Abelian magnetic flux through the 'hole'. Matter fields going around the hole will be transformed by a phase factor, just as in the Bohm-Aharonov effect in quantum electrodynamics. The unbroken gauge group which survives down to low energies in four dimensions will be that which is left invariant by all the phase factors associated with these holes. If one compactifies on a Calabi-Yau manifold, this surviving gauge group contains at least one extra gauge boson beyond those present in the Standard Model, whose phenomenological signatures we will discuss below.

Topology plays yet another important role in the effective theory, by constraining severely the self-couplings of the matter fields²³. Some of these couplings may even vanish for topological reasons, possibly providing a clue to the patterns of quark and lepton masses. No wonder particle phenomenologists are reaching anxiously for their mathematics textbooks! Topology may provide the answer to the flavour problem.

The effective four-dimensional theory arrived at after all these topological machinations is a simple supergravity theory with a rather special property: its field energy is completely independent of the values of some of the fields. Such no-scale supergravity theories had been advocated previously^{24,25} as a framework in which it might be possible to determine dynamically the mass scale of the weak interactions. The idea was that the flatness of the effective potential in certain field directions would be broken by quantum corrections, which would generate an energetically preferred value for the previously undetermined fields. In this way one would hope to derive the hierarchy $m_W/m_P \ll 1$, and thus finally solve the mass problem.

For this model to work, one needs supersymmetry to be broken, and so far we have been safeguarding it jealously. A popular idea is that supersymmetry is broken spontaneously by strong interactions in the second, hidden, E_6' gauge group, which cause the supersymmetric fermion partners of the hidden gauge bosons to condense in the vacuum^{26,27}, much like quarks in QCD and Cooper pairs in superconductivity. Even if this does occur, it is not clear how the scale of supersymmetry breaking is determined nor how this is related to the masses of the physical 'observable' sparticles—squarks ($\tilde{q}$), sleptons ($\tilde{l}$), gluinos ($\tilde{g}$), photinos ($\tilde{\gamma}$), and so on. Be that as it may, the effective low-energy theory will have the general structure of a supersymmetric gauge theory supplemented by small supersymmetry-breaking masses for these as yet unseen sparticles.

Towards experimental tests

Although many aspects of the above model are unclear and dubious², this has not prevented some optimistic souls from studying the possible low-energy signatures of the superstring which it suggests. As was mentioned above, Calabi-Yau compactification leaves at least one additional neutral gauge boson beyond those in the Standard Model, although it may be absent in other models for compactification. The fact that no extra neutral gauge boson has been seen along with the conventional Z^0 at the CERN $p\bar{p}$ collider means that it must weigh at least 110 GeV^{28,29}. However, its indirect effects could be detectable

even before the new boson itself. For example, in general it mixes with the Standard Model Z^0 and decreases its mass. The agreement of the observed Z^0 mass with the value predicted in the Standard Model suggests that the additional neutral gauge boson should weigh at least 200 GeV³⁰. This lower bound is more stringent than that imposed by the agreement of low-energy neutral-current measurements with the predictions of the Standard Model^{28,29,31}. Detailed measurements of the properties of the Z^0 in e^+e^- collisions at the SLAC (Stanford Linear Accelerator Center) linear collider or at LEP offer the best prospects for advancing our knowledge in this area.

Turning now to the extra matter fields in the superstring at low energies, one possible problem is provided by the right-handed neutrinos. Unless they are either much heavier than the conventional left-handed neutrinos, or else very weakly coupled to the more conventional matter particles, they could change the rate of expansion of the Universe during primordial nucleosynthesis and spoil the successes of the classic calculations of light element abundances³². As it is unclear how to give large masses to the right-handed neutrinos, people have tried to estimate how weakly coupled they need to be, which translates into a cosmological lower bound on the mass of the extra gauge boson. There is some debate as to just how heavy it should be^{32,33}, but a lower bound of several hundred GeV seems likely. This is not necessarily an embarrassment for model-builders, although it may delay the discovery of new superstring particles.

One of the clearest low-energy signatures of the superstring may be provided by the additional charge $-1/3$ 'quarks' appearing in each E_6 generation. The word 'quark' is put in quotation marks because its interactions cannot be the same as those of conventional quarks. Invariance under E_6 requires it to couple directly to either a pair of conventional antiquarks, or a quark and a lepton. It cannot have both couplings simultaneously, or baryon and lepton numbers would be violated and the proton would decay very rapidly. Just how the superstring avoids this disaster is another mystery, but assuming that it does, one has a novel signature for the superstring. The new scalar 'quarks' may decay either into pairs of conventional quarks, detectable by experiment as a pair of hadronic jets, or into a lepton and a quark, detectable as a lepton plus a jet of hadrons.

Superstring phenomenology is still a very young subject, and it is doubtless premature to place any faith in details of the general model outlined above. There are many open questions and technical problems, and it is easy to ridicule superstring advocates for their totalitarian fervour. However, in the words of a candy wrapper I opened a few years ago: "It is only the optimists who achieve anything in this world". Even if many features of the above model are wrong, new ideas are being brought into particle physics at a rate unequalled since the renaissance of gauge theory in 1971. Our intellects are being mathematically stimulated, and we are thinking of many new types of phenomena that our experimental colleagues can search for. We cannot discover the secrets of nature by pure reason, and must look to an experimental breakthrough. At the very least, the superstring may point out to us a previously unremarked stone which, when turned over, may reveal interesting new life beneath.

1. Green, M. B. *Nature* **314**, 409-414 (1985).
2. De Rújula, A. *Nature* **320**, 678 (1986).
3. Llewellyn-Smith, C. H. *Nature* **312**, 588-592 (1984).
4. Ellis, J. & Peccei, R. (eds) CERN Rep. No. 86-02 (1986).
5. Ellis, J. *Nature* **313**, 626-627 (1985).
6. Chamseddine, A. H. *Nucl. Phys.* **B185**, 403-415 (1981).
7. Bergshoeff, F., de Roo, M., de Wit, B. & van Nieuwenhuizen, P. *Nucl. Phys.* **B195**, 97-136 (1982).
8. Chapline, G. & Manton, N. *Phys. Lett.* **120B**, 105-109 (1983).
9. Green, M. B. & Schwarz, J. H. *Phys. Lett.* **149B**, 117-122 (1984); **151B**, 21-25 (1985).
10. Lovelace, C. *Phys. Lett.* **34B**, 500-506 (1971).
11. Schwarz, J. H. *Nucl. Phys.* **B46**, 61-74 (1972).
12. Neveu, A. & Scherk, J. *Nucl. Phys.* **B36**, 155-161 (1972).
13. Yoneya, T. *Prog. Theor. Phys.* **51**, 1907-1920 (1974).
14. Scherk, J., Schwarz, J. H. *Nucl. Phys.* **B81**, 118-144 (1974).

15. Candelas, P., Horowitz, G. T., Strominger, A. & Witten, E. *Nucl. Phys.* **B258**, 46-74 (1985).
16. Calabi, E. in *Algebraic Geometry and Topology: a Symposium in Honor of S. Lefschetz* 78-89 (Princeton University Press, 1957).
17. Yau, S. T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1798-1799 (1974).
18. Witten, E. preprint, Princeton Univ. (1985).
19. Witten, E. *Nucl. Phys.* **B258**, 75-100 (1985).
20. Dine, M., Kaplunovsky, V., Mangano, M., Nappi, C. & Seiberg, N. *Nucl. Phys.* **B259**, 549-571 (1985).
21. Breit, J. D., Ovrut, B. A. & Segrè, G. *Phys. Lett.* **1** **158B**, 33-39 (1985).
22. Hosotani, Y. *Phys. Lett.* **126B**, 309-313 (1983).
23. Strominger, A. & Witten, E. *Commun. Math. Phys.* **101**, 341 (1985).
24. Ellis, J., Lahanas, A. B., Nanopoulos, D. V. & Tamvakis, K. *Phys. Lett.* **134B**, 429-435 (1984).
25. Ellis, J., Kounnas, C. & Nanopoulos, D. V. *Nucl. Phys.* **B241**, 406-428; **B247**, 373-395 (1984).
26. Derendinger, J.-P., Ibañez, L. & Nilles, H.-P. *Phys. Lett.* **155B**, 65 (1985).
27. Dine, M., Rohm, R., Seiberg, N. & Witten, E. *Phys. Lett.* **156B**, 55-60 (1985).
28. Cohen, E., Ellis, J., Enqvist, K. & Nanopoulos, D. V. *Phys. Lett.* **165B**, 76-82 (1985).
29. Barger, V., Deshpande, N. G. & Whisnant, K. *Phys. Rev. Lett.* **56**, 30-33 (1986).
30. Ellis, J., Enqvist, K., Nanopoulos, D. V. & Zwirner, Nucl. Phys. **B276**, 14-70 (1986); *Mod. Phys. Lett.* **A1**, 57-69 (1986).
31. Durkin, L. S. & Langacker, P. *Phys. Lett.* **166B**, 436-442 (1986).
32. Ellis, J., Enqvist, K., Nanopoulos, D. V. & Sarkar, S. *Phys. Lett.* **167B**, 457-463 (1986).
33. Steigman, G. A., Olive, K. A., Schramm, D. N. & Turner, M. S. *Phys. Lett.* **176B**, 33-38 (1986).

ARTICLES

Non-explosive silicic volcanism

J. C. Eichelberger, C. R. Carrigan, H. R. Westrich & R. H. Price

Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

Silicic magma can erupt quietly, as vapour-poor lava, despite petrological evidence that it once contained ample dissolved water to drive violent venting of tephra. Non-explosive eruption of lava appears to result from rapid, sub-surface gas release from magma ascending as a permeable foam. The degassed foam then collapses during extrusion. Conditions of shallow ascent, rather than pre-eruption volatile concentrations, control eruptive behaviour.

SILICIC volcanic eruption sequences commonly begin with vigorous ejection of fragmental magmatic material consisting of pumice and ash, and conclude with quiet effusion of lava domes or flows. The explosive case is well described by considering exsolution of volatiles in magmatic melt during decompression, expansion of the magma as a foam, and subsequent fragmentation and further expansion of the magma as a dusty gas^{1,2}. No such theoretical treatment has been offered for the non-explosive case. The absence of fragmentation during extrusion clearly reflects a lower vapour content, but this may arise in two ways. If the magma behaves as a chemically closed system, then the extrusion of domes results from ascent of initially volatile-poor magma, and the transition from explosive to non-explosive behaviour reflects pre-eruption volatile gradients in the parent body³. This is a widely held view of the relationship between eruptive behaviour and pre-eruption magma composition. Alternatively, if the magma behaves as an open system during ascent, then the transition to non-explosive behaviour reflects a change in the conditions of ascent. Here we present a gas-dynamical model for dome-forming eruptions, involving open-system behaviour of magma on an eruptive timescale. Such a model necessarily proceeds from an accurate physical and chemical description of an appropriate igneous system. Research drilling into the 600-yr-old Obsidian Dome⁴, the largest dome of the Inyo Domes chain in eastern California, has recently provided information that could not be obtained from examining the surfaces of young domes and the intrusive roots of old ones. Although we base our discussion in part on results from this system, we believe that the model has general applicability.

Obsidian Dome

The physical characteristics of Obsidian Dome and the upper 600 m of its intrusive feeder are interpreted from surface observations and three continuously cored holes (Fig. 1). The first hole penetrated the distal portion of the dome, the second passed through the proximal section of the dome and the underlying intrusion that forms the dome's conduit, and the third intersected an unvented portion of the same intrusion, ~1 km to the south.

Obsidian Dome is pancake-shaped, 0.1 km³ in volume, 2 km in diameter, and 30-50 m thick at its margins. Near the vent, the thickness increases rapidly to 100 m. At a depth of 400-500 m, the conduit of the dome is 33 m wide. The conduit lies

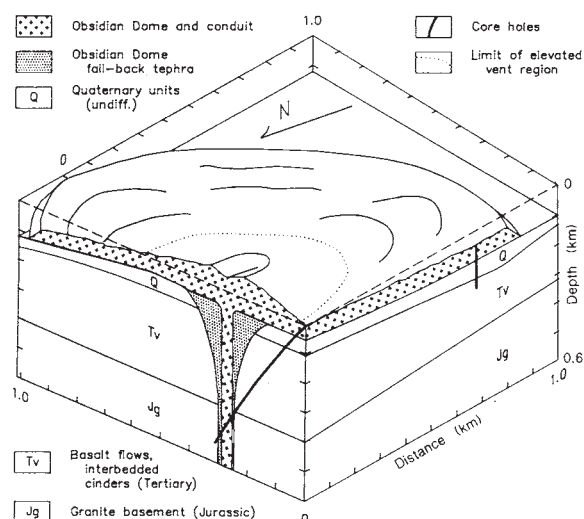


Fig. 1 Block diagram showing locations of the distal and proximal drill holes at Obsidian Dome and major intersected subsurface features. The origin of the plot is the wellhead of the proximal hole, which intersects the conduit of the dome; the two vertical planes contain the holes and form a 99° angle. A third hole, not shown, lies approximately on the right rear face of the block, and intersects an unvented portion of the feeder dike at 610-640 m depth.

within a broader vent structure, exceeding 50 m in width at the depth of observation, which contains fall-back tephra, slumped blocks of wall rock, and small co-magmatic intrusions. These materials are presumed to occupy the lower portion of a vent funnel which was excavated during the immediately preceding explosive phase of activity. One kilometre to the south, the intrusion is 7 m wide at a depth of 600 m, where it is in direct contact with Sierra Nevada granitic basement. The intrusion is evidently a long, thin dike which widens beneath the dome.

There are no apparent petrological differences between the explosively erupted tephra and non-explosively erupted dome,

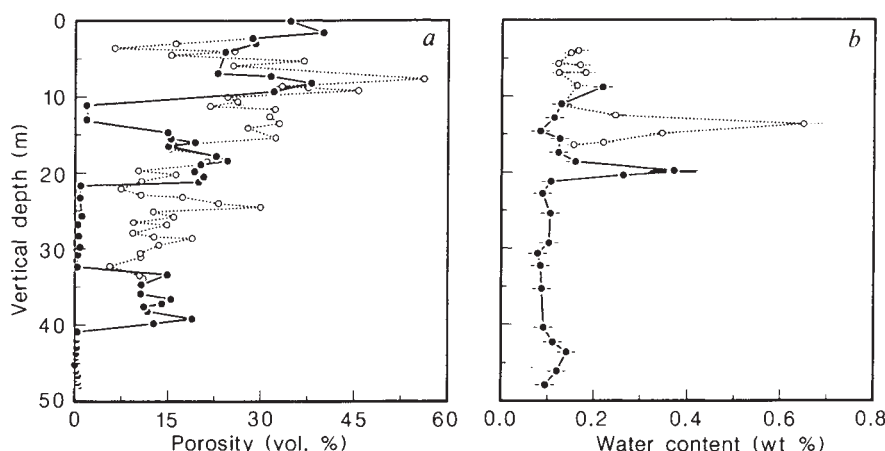


Fig. 2 Porosity (a) and water content (b) profiles through the distal (●) and proximal (○) portions of Obsidian Dome. a, Porosities were calculated from bulk and grain density measurements; replicate determinations gave errors (2 s.d.) of ± 0.3 vol. %. Porosity attributed to second boiling following emplacement (consisting of large voids associated with spheroidal regions of devitrification-lithophysae) is not included in the profiles. b, Water contents were determined for glassy samples by Karl Fischer titration. Error bars show 2 s.d., where larger than symbols. Peaks in water content near the top of both sections appear to reflect the adsorption on bubble walls at sub-magmatic temperature of magmatic water vapour escaping from the flow interior, because these samples are from bubble-rich zones and release water at relatively low temperature.

although differences would be expected if volatile concentrations varied greatly in the parent magma reservoir. The rocks range in composition from rhyolite to rhyodacite. Phenocrysts are <10% by volume and include the hydrous phases hornblende and biotite. In the dome, glass-rich zones extend 10–20 m inward from the top and bottom surfaces to a more crystalline central portion. In contrast, samples from the intrusion are completely crystalline, except for one 10-cm-thick glassy contact zone.

Porosity data (Fig. 2a) provide evidence of the final distribution of vapour in the system. Like all silicic domes, the surface of Obsidian Dome is pumiceous. At the dome margin, an obsidian (bubble-free glass) layer is exposed ~10 m below the flow top. Observations from the distal core hole show that obsidian comprises a large portion of the distal interior. However, obsidian is absent near the vent, where all the lava is vesicular. Similarly, no bubble-free material occurs in the intrusion, where most porosity values lie between 10 and 20%.

Water is the predominant volatile component in the glasses, and therefore the most important chemical component in determining eruptive behaviour. Figure 2b shows that the low values previously observed at the surface of the dome⁵ continue into the interior. Of special interest is the water content of the obsidian, as this material could not have changed chemically since emplacement. The retained magmatic water content of obsidian in the dome is everywhere close to 0.1 wt %. In contrast, the variable water contents of pumiceous lava probably reflect adsorption during cooling. None of the glass within the intrusion is sufficiently unaltered to allow the determination of magmatic water values. However, fresh, intrusion-derived glass fragments in nearby fractures yielded water contents of 1.15 and 1.21 wt % at depths of 290 and 370 m, respectively.

Compositional path of ascending magma

The above observations can be used to determine whether the water content of the dome-forming magma changed during ascent. The pre-eruption water content is constrained by the phenocryst assemblage, the surface-emplacement water content can be directly determined from the obsidian, and inferences about water content at intervening depths can be drawn from characteristics of the intrusion. Experiments⁶ in a synthetic system similar to Obsidian Dome show the dependence of liquidus temperature and hornblende stability on water content at 2 and 8 kbar. Liquidus temperature declines steeply from >1,200 °C in the anhydrous system to 1,000 °C at 4 wt % water. Hornblende becomes stable at water contents of 3–4 wt %. Higher pressure does not appreciably affect the liquidus or the minimum water content at which hornblende is stable. The low phenocryst content indicates that the magma was near its liquidus at ~950 °C (as determined by Fe–Ti oxide thermometry;

ref. 7). The phenocryst assemblage thus requires ≥ 3 wt % water in melt before ascent.

Figure 3 shows what happens to magma with 3 wt % water as it approaches the surface. We can assume that the ascending magma is under a confining pressure close to lithostatic because of the fractured nature of the intruded environment. The condition for stability of vapour bubbles in the magma is that the vapour pressure in the bubbles equals the confining pressure on the magma. Bubble growth occurs in response to excess vapour pressure, but the required excess in hydrous magma is believed to be small¹. Bubble content increases with decreasing depth because confining pressure decreases, causing water to become less soluble in melt and the vapour to expand. For the system of interest, bubble growth begins at 500 bar, or ~2.5 km depth, and the bubble:melt ratio reaches 3:1 at ~400 m depth. This degree of inflation is important because significantly higher porosities are seldom observed in ejecta¹. The stretched and weakened bubble walls apparently fail at this point, and an explosive eruption follows. The dome magma did not reach fragmentation conditions, although its equilibrium bubble:melt ratio at 3 wt % water would have been 400:1 on reaching the surface or, alternatively, its excess pore pressure at observed porosities and 3 wt % water would have been 100–500 bar.

If we consider the final composition of the dome magma as represented by the obsidian, we find that the water content is consistent with an absence of bubbles in melt under the existing loads of 2–10 bar. However, it is apparent from the above discussion that the phase assemblage must be regarded as relict from a much more water-rich system. The liquidus temperature at 0.1 wt % would exceed 1,200 °C, in conflict with Fe–Ti oxide compositions, and hornblende would not be stable at any temperature. Instead, the observed water content suggests open-system equilibration with surface conditions, because it corresponds closely with the expected solubility of water in melt at atmospheric pressure⁸. Furthermore, the hydrogen isotopic composition of the water lies at the end-point of a trend defined by more water-rich explosive ejecta, which is best explained by open-system Rayleigh fractionation of hydrogen isotopes between melt and vapour during the release of water⁵.

The loss of water during ascent is also indicated by two features of the intrusion. First, the boundary glass zones are more than two orders of magnitude thicker in the dome than in the intrusion, and are entirely absent from the conduit. This cannot be explained by thermal history because the relative sizes indicate that the intrusion cooled as fast as, or faster than, the dome⁹. As water facilitates crystallization in rhyolitic systems¹⁰, greater retention of water at depth would explain the contrast in crystallinity between the dome and its feeder. Second, vesicles that appear to be primary occur throughout the intrusion. Their existence as water vapour bubbles would require at least 1 wt %

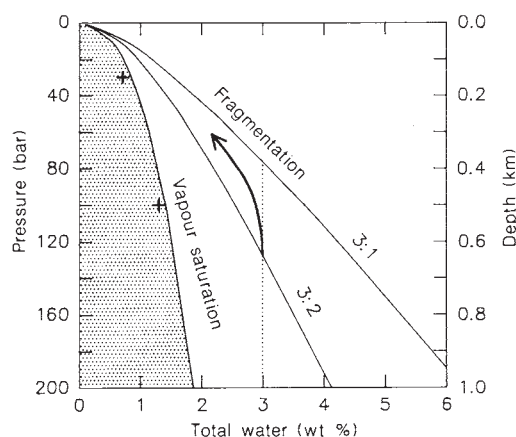


Fig. 3 Total magmatic water content for crystal-poor rhyolite magma as a function of pressure and depth in the uppermost 1 km. Ideal gas behaviour of water and solubility relations as described by Burnham¹⁹ are assumed. The depth dependence involves the additional assumptions of pressure equilibrium between lithostatic load and vapour pressure, and an average bulk density of overburden of 2 g cm^{-3} . The vapour-undersaturated region is shaded. Two values (crosses) from the solubility approximation (equation (3)) used by Sparks¹ are shown; this approximation is used in subsequent calculations. Two curves describe the system at bubble:melt ratios of 3:2 (60% porosity) and 3:1 (75% porosity). The latter curve is believed to mark the onset of fragmentation. Magma that behaves as a chemically closed system during ascent will follow a vertical path (dotted line) and fragment to form tephra. The inferred behaviour of dome magma is indicated by the arrow.

water in the coexisting melt (Fig. 3), consistent with the 1.2 wt % water observed in the glass fragments. If formed secondarily during crystallization (second boiling), more water than is present in the obsidian would be required to produce the observed porosity.

These considerations require that the dome magma lost most of its water during ascent, and that the magma extruded as Obsidian Dome lost more water than the magma emplaced at 400–600 m depth.

Physical conditions of water release

Observations at Obsidian Dome also provide evidence of the process by which water release occurs, and the time and distance scales over which the process operates. Degassing must occur on an eruptive timescale, otherwise the magma would fragment during ascent. The timescale of ascent and hence decompression is severely constrained by the small volume of the feeder relative to the dome. If the dome was extruded in a year or less¹¹, and the conduit is twice as long along the dike trend as its intersected width, then the dome magma underwent the final 100 bar of decompression in two days or less. Magma in the form represented by obsidian could not change chemically on this timescale, because mass transport would have to occur exclusively by the exceedingly slow process of diffusion in melt¹². However, the increasing porosity of the dome as the vent is approached (Fig. 2a) suggests that the magma was extruded as a foam that collapsed as it flowed outwards. Indeed, the ascent of magma as a foam is required by the inferred initial water content. If the bubbles of the foam were connected, then volatiles could escape the system by a two-step process. The first step involves diffusion of volatiles through melt into bubbles, and the second the flow of gas from bubble to bubble and out of the system. The first step occurs over micrometres and requires seconds¹. The second occurs over metres, and calculation of the timescale requires knowledge of the permeability of the magmatic foam.

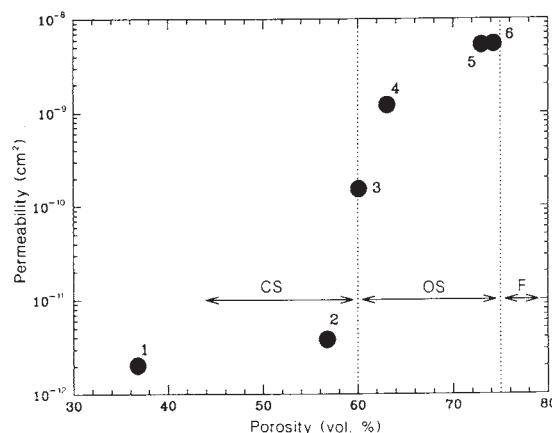


Fig. 4 Gas permeability and porosity for Inyo rhyolite and related samples. Gas permeabilities were determined with nitrogen at room temperature, using confining pressures of 3–50 bar and input pressures of 1.2–50 bar. Less permeable samples required higher pressures. The confining pressure was varied to check for leakage between the sample and the jacket. The range of permeabilities over a variety of flow conditions is nowhere greater than the symbol size. Samples are from: 1, proximal hole, 5.3 m depth; 2, proximal hole, 14.3 m depth; 3, proximal hole, 13.9 m depth; 4, 5, Panum Dome, surface; 6, Large pumice block, explosive phase of Inyo eruption (Glass Creek vent), collected 1 km south-west of Obsidian Dome. Porosity regimes: CS, closed system; OS, open system; F fragmentation.

Once out of the magma, gas could easily escape into the 'sand-pile' environment of finely fragmented material provided by the vent funnel.

No direct measurements of permeability of vesicular, non-fragmental magmatic material have previously been made, although Witham and Sparks¹³ inferred high permeability from the behaviour of pumice in water. Figure 4 shows the gas permeability of Obsidian Dome and related rhyolite over the porosity range 37–74%. The permeability increases by three orders of magnitude during the expansion of bubbles from 55 to 75 vol. % of the system. Hence, the bubbles become well interconnected during ascent before the fragmentation regime is reached—at depths of $\geq 500 \text{ m}$ for magma with the inferred water content of $\geq 3 \text{ wt } \%$ (Fig. 3).

Porous-flow model

The escape of gas from a foam can be modelled using a combined analytical and numerical approach. The ascending magma is treated as a dike, with outward horizontal gas flow symmetrical about its centre-line, where the flux is zero. A constant, zero gas pressure is applied at the margins, to simulate the large discontinuity in permeability which exists between the porous magma, which provides some resistance to gas flow, and the more permeable fragmental material of the vent funnel, into which the magma has intruded. The gas flow is described by Darcy's law¹⁴, which relates the pore velocity u to the pressure gradient dP/dx , through the porosity ϕ , the permeability K , and the gas viscosity μ :

$$\phi u = -\frac{K}{\mu} \frac{dP}{dx} \quad (1)$$

The amount of gas vapour in an elementary volume is given by a conservation equation:

$$\phi \frac{\partial \rho_v}{\partial t} = -\phi \frac{\partial}{\partial x} \rho_v u + \rho_m (1 - \phi) \frac{d\chi}{dP} \frac{\partial P}{\partial t} \quad (2)$$

where ρ_v and ρ_m are the vapour and melt densities, respectively,

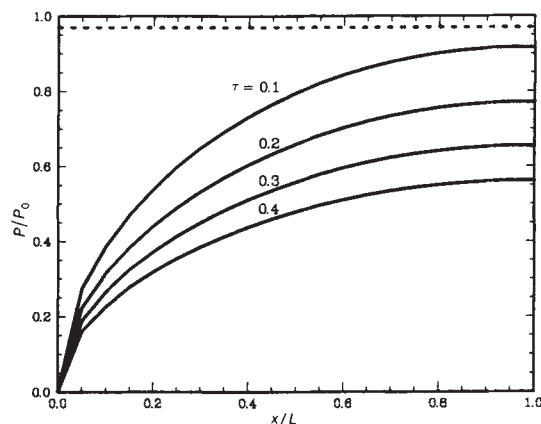


Fig. 5 Calculated vapour pressure profiles in magma across the half-width of a dike, at various times (τ) after the onset of degassing. Vapour pressure, position and time are expressed as dimensionless variables. Distance increases from the edge of the dike ($x/L=0$) to the centre-line ($x/L=1.0$). The scalings are: P_0 , the initial vapour pressure; L , the half-width of the dike; and $\mu L^2 \phi / k P_0$, the gas diffusion timescale. Note that time scales with the square of the dike half-width. For Obsidian Dome we use $P_0=100$ bar, $L=1,650$ cm, $\mu=4.3 \times 10^{-4}$ poise (water vapour at 900°C), $\phi=0.67$, and $K=0.2 \times 10^{-8}$ cm². Thus, to obtain the Obsidian Dome case, multiply (P/P_0) by 100 to obtain bars, (x/L) by 16.5 to obtain metres, and τ by 2,600 to obtain seconds. The four profiles are then at 375, 750, 1,125 and 1,500 s. At 1,500 s, the load on the magma ascending at 1 cm s^{-1} has decreased by only 3 bar (indicated by the dashed line), but the vapour pressure has decreased by >40 bar.

t is time, and χ is the weight fraction of water in the melt. Equation (2) relates the time rate of change of the local vapour density to the divergence of the mass flow rate, and a pressure-dependent vapour source term which takes into account the exsolution of volatiles from the melt. Sparks' approximation¹ for the relationship between equilibrium melt water content and vapour pressure,

$$\chi = aP^{1/2} \quad (3)$$

where $a=0.0013 \text{ bar}^{-1/2}$, permits the evaluation of the vapour source term. Finally, the ideal gas law provides a satisfactory description of water vapour expansion². Combining equations (1)–(3) and the ideal gas law, we obtain a dimensionless equation for vapour pressure as a function of time:

$$\frac{\partial P}{\partial t} = \frac{P^{1/2}}{P^{1/2} - C} \frac{\partial}{\partial x} P \frac{\partial P}{\partial x} \quad (4)$$

where $C = [\rho_m(1-\phi)RTa]/2\phi P_0^{1/2}$, R is the gas constant, T is temperature, and P_0 is the vapour pressure at $t=0$. This equation may be solved numerically over the half-width of the dike for the stated boundary conditions (Fig. 5).

To apply the solution to a given eruptive event, the relevant physical parameters are needed to evaluate the scalings for pressure, time and horizontal coordinate. If we assume that degassing begins as magma enters the base of the vent funnel and that this depth at Obsidian Dome is 500 m, then the bubble:melt ratio of the magma will be $\geq 2:1$ (Fig. 3) and the permeability will be $\geq 0.2 \times 10^{-8} \text{ cm}^2$ (Fig. 4). These values lie between the opening of the system to gas flow and fragmentation. The use of this value of porosity implies that significant bubble collapse occurred following degassing, particularly in the magma comprising the conduit and distal portion of the dome, but this is consistent with the central depression in the vent region and rapid decrease in dome thickness away from the vent region (Fig. 1). The half-width used, 16.5 m, is derived from the drilling results. Other values of physical properties used in the

scalings are typical for rhyolitic magma (for example refs 1 and 12).

We can now ask whether degassing is sufficiently rapid to prevent further bubble growth, which would lead to fragmentation. Bubble growth will not occur if the vapour pressure declines as fast or faster than the lithostatic load on the ascending magma. A reasonable upper bound for the ascent rate is 1 cm s^{-1} (dome formation in 2 months), corresponding to a decompression rate of $\sim 0.002 \text{ bar s}^{-1}$. The calculation shows that pore pressure declines at 0.03 bar s^{-1} , an order of magnitude faster (Fig. 5). Although we have simplified the problem by keeping porosity and permeability constant, the calculation demonstrates that if the magma ascends through an environment in which gas can escape freely, it cannot reach the degree of inflation observed in the explosive ejecta. Instead, the magma inflates until the contributions to bubble volume from vapour exsolution and expansion are balanced by increasingly rapid vapour loss, and arrives at the surface fully degassed (Fig. 3). Vertical gas flow may predominate late in the ascent.

The uniform 0.1 wt % water content of obsidian in the dome shows that the extruding foam was rather stiff. The vapour pressure throughout the dome bled off to near-atmospheric pressure before bubble collapse began. This left the system water-undersaturated when pressure equilibrium was re-established, so that the small amount of vapour remaining, 0.02 wt % at 50% porosity, was readily resorbed to produce bubble-free obsidian. Had bubble collapse restored the equilibrium between lithostatic load and pore pressure before degassing was complete, then the water content under the 10-bar load near the base of the dome would be measurably higher ($\sim 0.5 \text{ wt %}$; Fig. 3). The pumiceous carapace did not undergo collapse because of rapid cooling and a load pressure of <2 bar. The process is analogous to welding in ash-flow tuffs, which can produce obsidians without bubbles or even relict bubble textures¹⁵, and with water contents identical to dome obsidians⁸. The heat required for welding¹⁶ is better conserved in a dome than in a pyroclastic deposit.

Discussion

Our model indicates that conditions during final ascent, when silicic magma is permeable to gas flow ($\leq 1 \text{ km}$ depth for the commonly postulated water contents of 3–5 wt %^{17,18}) will determine whether an explosive eruption occurs. The development of a tephra-filled vent funnel during initial explosive events provides the high-permeability environment necessary for subsequent lava extrusion. Thus, a decrease in the explosivity of events does not demonstrate volatile zonation in the deep parent reservoir. The high permeability of vent funnels may be short-lived, however, because the fall-back tephra will itself tend to weld. Thus, activity that does not quickly follow funnel formation may revert to vent-clearing explosions. The fact that all large-volume silicic eruptions are explosive may reflect the exponential increase in degassing time with width of the ascending body (Fig. 5). The largest known rhyolitic lava flows have volumes of $<10^2 \text{ km}^3$ (ref. 20), whereas the volumes of individual explosive rhyolitic eruptions range to $>10^3 \text{ km}^3$ (ref. 21). It appears likely that magma in such large eruptions, rising through conduits of perhaps $\geq 100 \text{ m}$ in width, cannot degas sufficiently during ascent to prevent explosive fragmentation.

This view of silicic volcanism solves the paradox posed by the contrasting vapour content but petrological similarity of co-magmatic tephra and dome eruptions^{22–24}. It should motivate a re-examination of the shallow intrusive environment in relation to eruptive behaviour.

We are grateful for help and suggestions from David McTigue, Melvin Baer, Mark Stavig, Shigeo Aramaki, Peter Lipman and Stephen Sparks. This work was supported at Sandia National Laboratories by the US Department of Energy under contract DE-ACO4-76DP00789.

Received 28 March; accepted 6 August 1986.

1. Sparks, R. S. J. *J. Volc. geotherm. Res.* **3**, 1-38 (1978).
2. Wilson, L. *J. Volc. geotherm. Res.* **8**, 297-313 (1980).
3. Fink, J. H. *Geol. Soc. Am. Bull.* **94**, 362-380 (1983).
4. Eichelberger, J. C., Lysne, P. C., Miller, C. D. & Younker, L. W. *Eos* **66**, 186-187 (1985).
5. Taylor, B. E., Eichelberger, J. C. & Westrich, H. R. *Nature* **306**, 541-545 (1983).
6. Nancy, M. T. *Am. J. Sci.* **283**, 993-1033 (1983).
7. Vogel, T. A., Schuratz, B. C. & Younker, L. W. *Eos* **66**, 384 (1985).
8. Eichelberger, J. C. & Westrich, H. R. *U.S. geol. Surv. Open File Rep. No. 84-939*, 147-150 (1984).
9. Kasameyer, P. W., Younker, L. W., Eichelberger, J. C., Lysne, P. C. & Vogel, T. A. *Eos* **66**, 385 (1985).
10. Tuttle, O. F. & Bowen, N. L. *Mem. geol. Soc. Am.* No. 74 (1958).
11. Macdonald, G. A. *Volcanoes* (Prentice-Hall, Englewood Cliffs, 1972).

12. Shaw, H. R. *Publs Carnegie Instn* **634**, 139-170 (1974).
13. Witham, A. G. & Sparks, R. S. *Bull. volc.* (in the press).
14. Bird, R. B., Stewart, W. E. & Lightfoot, E. N. *Transport Phenomena* (Wiley, New York, 1960).
15. Ross, C. S. & Smith, R. L. *Prof. Pap. U.S. geol. Surv.* **366**, 1-77 (1961).
16. Riehle, J. R. *Geol. Soc. Am. Bull.* **84**, 2193-2216 (1973).
17. Whitney, J. A. *J. geol.* **83**, 1-31 (1975).
18. Ritchey, J. L. *J. Volc. geotherm. Res.* **7**, 373-386 (1980).
19. Burnham, C. W. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 71-136 (Wiley, New York, 1979).
20. Christiansen, R. L. *U.S. Geol. Surv. Open File Rep. No. 84-939*, 743-783 (1984).
21. Lipman, P. W. *J. geophys. Res.* **89**, 8801-8841 (1984).
22. Lipman, P. W. *J. Geol.* **79**, 438-456 (1971).
23. Bacon, C. R. *J. Volc. Geotherm. Res.* **18**, 57-115 (1983).
24. Ewart, A., Hildreth, W. & Carmichael, I. S. E. *Contr. Miner. Petrol.* **51**, 1-27 (1975).

LETTERS TO NATURE

Variable radio source GT0116+622 is a possible counterpart to Cas γ -1

P. C. Gregory*, N. Duric*, A. Reid*, J. Picha*,
T. Stevenson* & A. R. Taylor†

* Department of Physics, University of British Columbia, Vancouver,
British Columbia, Canada V6T 1W5

† Kapteyn Astronomical Institute, University of Groningen,
9700 AV Groningen, The Netherlands

Identification of celestial γ -ray sources has proved to be very difficult because of the large positional uncertainties (typically several degrees) and the transient nature of many of these objects. Highly variable radio emission is a striking characteristic of the most secure source of very high-energy radiation, Cygnus X-3 (ref. 1). Here we report the discovery of another highly variable radio source which coincides in position with the transient TeV γ -ray source Cas γ -1. We have obtained an accurate radio position with the Very Large Array (VLA) telescope but no optical counterpart was found on the Mount Palomar Sky Survey prints. We have also obtained upper limits on radio emission from the binary X-ray pulsar, 4U0115+634, another proposed identification for Cas γ -1.

The variable radio source GT0116+622 was discovered during the University of British Columbia (UBC) galactic radio patrol project²⁻⁴ carried out between 1977 and 1984 with the National Radio Astronomical Observatory (NRAO) 91-m telescope at a wavelength of 6 cm. The aim of this project was to survey repeatedly a strip along the galactic plane ($l = 40^\circ$ to

220° , $b \leq \pm 2^\circ$) for transient radio sources. In the survey area of ~ 500 square degrees, 32 variables and 27 possible variables were detected. Cyg X-3 was the only previously known source in this sample. Short term variables make up 64% of the sample (timescale of a few days) and among these, GT0116+622 exhibits the fourth largest degree of variability following Cyg X-3, GTO351+543a and GT0236+610. The probability of a chance detection of a short-term variable radio source comparable to GT0116+622 in the $1^\circ \times 1^\circ$ γ -ray error box is ~ 0.008 .

Table 1 lists 53 daily flux density measurements of the source between 1977 and 1984. Over this period the flux density varied from ≤ 5 to 73 mJy. The extensive coverage of the observations in 1977 and 1984 shows significant variability on timescales as short as 1 day with some evidence for a characteristic timescale of 5 days.

Ten-minute photographic VLA observations in A configuration were obtained on 1 January 1985 at 6 and 20 cm. A single unresolved source was detected at both wavelengths within the positional uncertainty of the radio patrol results. The observations were self-calibrated and the resulting maps noise-limited at both wavelengths to a dynamic range of $\sim 200:1$. The measured flux densities are: 20-cm flux density = 57.0 ± 0.5 mJy; 6-cm flux density = 40.5 ± 0.5 mJy. The 6-cm VLA flux density corresponds to about the middle of the range observed in the patrol measurements. The spectral index between 20 and 6 cm of $\alpha = -0.3$ is typical of many compact synchrotron-emitting variable radio sources, both galactic and extragalactic.

Table 2 compares our VLA position for GT0116+622 with the best position for Cas γ -1 obtained by Stephanian *et al.*⁵. We also include the position of a nearby binary X-ray pulsar, 4U0115+634, suggested⁶ as a possible counterpart to Cas γ -1. The position for 4U0115+634 is taken from ref. 7.

Table 1 Daily flux density measurements for GT0116+622

Date	Flux (mJy)	Date	Flux (mJy)	Date	Flux (mJy)	Date	Flux (mJy)
14 August 1977	25 $\pm$ 5	14 August 1978	68 $\pm$ 6	19 August 1981	30 $\pm$ 5	6 August 1984	39 $\pm$ 5
15 August 1977	19 $\pm$ 5	19 August 1978	53 $\pm$ 6	25 August 1981	35 $\pm$ 5	7 August 1984	42 $\pm$ 5
16 August 1977	15 $\pm$ 5	21 August 1978	68 $\pm$ 6	31 August 1981	31 $\pm$ 5	8 August 1984	30 $\pm$ 5
17 August 1977	10 $\pm$ 5	24 August 1978	73 $\pm$ 6			9 August 1984	24 $\pm$ 5
18 August 1977	12 $\pm$ 5			19 July 1984	38 $\pm$ 5	10 August 1984	31 $\pm$ 5
20 August 1977	29 $\pm$ 5	3 August 1979	36 $\pm$ 6	21 July 1984	17 $\pm$ 5	11 August 1984	42 $\pm$ 5
21 August 1977	30 $\pm$ 5	6 August 1979	32 $\pm$ 6	22 July 1984	23 $\pm$ 5	12 August 1984	39 $\pm$ 5
22 August 1977	16 $\pm$ 5			23 July 1984	37 $\pm$ 5	13 August 1984	26 $\pm$ 5
23 August 1977	30 $\pm$ 5	23 August 1980	26 $\pm$ 6	24 July 1984	29 $\pm$ 5	14 August 1984	20 $\pm$ 5
24 August 1977	15 $\pm$ 5	28 August 1980	30 $\pm$ 6	25 July 1984	25 $\pm$ 5	15 August 1984	39 $\pm$ 5
25 August 1977	7 $\pm$ 5	2 September 1980	49 $\pm$ 6	27 July 1984	25 $\pm$ 5	16 August 1984	29 $\pm$ 5
26 August 1977	1 $\pm$ 5	7 September 1980	32 $\pm$ 6	28 July 1984	24 $\pm$ 5		
27 August 1977	6 $\pm$ 5	11 September 1980	46 $\pm$ 6	29 July 1984	38 $\pm$ 5		
28 August 1977	23 $\pm$ 5			30 July 1984	37 $\pm$ 5		
29 August 1977	21 $\pm$ 5			4 August 1984	38 $\pm$ 5		
30 August 1977	18 $\pm$ 5			5 August 1984	38 $\pm$ 5		

Table 2 Comparison of positions of 6T0116+622, Cas γ -1 and 4U0115+634

Source	Ra (1950)	Dec (1950)
GT0116+622	01 h 16 m 01.28 s $\pm$ 0.01 s	62°13'30.49" $\pm$ 1"
Cas γ -1	01 h 16 m $\pm$ 4 m	62° $\pm$ 1°
4U0115+634	01 h 15 m 13.8 s $\pm$ 2.9 s	63°28'38" $\pm$ 20"

The positional measurements favour an association between GT0116+622 and Cas γ -1. However, the position of Cas γ -1 appears to be based on drift scans at a fixed declination of 62° using an instrument with a geometrical half-angle of the field of view of 0.9°. Allowing for the typical Cerenkov image size of half-width 0.5° means that a positional coincidence⁶ with 4U0115+634 is not excluded. No optical counterpart to GT0116+622 was found to the limits of the National Geographic and Mount Palomar Observatory Sky Survey prints.

The radio patrol results also provide upper limits on any radio emission from 4U0115+634 on a total of eighteen days in 1980 and 1981. No compact radio source was detected on any occasion above a 3 σ upper limit which varied between 15 and 40 mJy.

Cas γ -1 was first detected in 1971 by Stepanian *et al.*⁸ at energies of 10¹¹–10¹² eV by means of Cerenkov light flashes from extensive air showers. Subsequent observations⁵ in 1972 and 1973 yielded a second detection in only 1 month out of 10 observed. The authors concluded that the source was variable and that during 'outburst' the detected flux was 3×10^{-10} photons cm⁻² s⁻¹. Lamb and Weekes⁶ have suggested that the binary X-ray pulsar 4U0115+634 (pulse period = 3.6 s) is the identification of Cas γ -1. This claim is based largely on positional proximity and the recent detection by Chadwick *et al.*⁹ of the 3.6-s pulse period at TeV energies.

Our detection of a new highly variable radio source, GT0116+622, close to the centre of the γ -ray error box provides an alternative identification for Cas γ -1. Furthermore, another of our highly variable radio sources, GT0236+610, is the most probable counterpart of the COS B γ -ray source CG135+1 (refs 10–14). Together with Cyg X-3, these results suggest that highly variable radio emission is a common manifestation of compact γ -ray sources.

We plan to obtain further radio observations of GT0116+622 and 4U0115+634 during the second phase of the galactic radio patrol project in 1986 using a new seven-feed 14-channel receiver system on the NRADO 91 m telescope. The 20-cm flux of GT0116+622 is large enough to attempt to determine the distance to the object using the 21-cm line. There is no Einstein or Exosat X-ray observation of the Cas γ -1 field and an X-ray counterpart to GT0116+622 would seem to be likely if the object is related to Cas γ -1.

This research was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. N.D. acknowledges support of an NSERC Fellowship. NRAO is operated by Associated Universities Inc., under contract with NSF.

Magnetic field corrections to solar oscillation frequencies

B. Roberts*‡ & W. R. Campbell†

* Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242, USA

† Department of Applied Mathematics, University of St Andrews, St Andrews, Fife, KY16 9SS, UK

‡ Present address: Department of Applied Mathematics, University of St Andrews, St Andrews, Fife, KY16 9SS, UK

The presence of a magnetic field both deep within the Sun and in its atmosphere raises the question of the field's influence on the p- and g-modes of oscillation and the implications for helioseismology. Observations^{1,2} of p-modes, in particular, have permitted a theoretical determination³ of the sound speed within the solar interior, thus providing a seismological probe of the Sun's depths. Magnetic fields within the Sun are likely to be too weak to significantly affect this determination of the sound speed. Nonetheless, magnetic fields may modify the oscillation frequencies in a distinctive fashion, thereby raising the possibility of placing limits on interior field strengths through frequency measurements⁴. Recently, Woodard and Noyes⁵ have reported a slight but systematic decrease in frequencies of low-degree p-modes from 1980 to 1984. Here we argue that the frequencies of both p- and g-modes are modified by a magnetic field. In particular, we attribute the decrease in p-mode frequencies to a magnetic field within the solar interior evolving over the solar cycle. Field strengths at the base of the convection zone of at least 5×10^5 G are required.

In the photospheric layers of the Sun, magnetic fields occur in concentrated flux-tube form, with field strengths ranging from 1.5 kG in the smallest tubes to 3 kG in sunspots^{6,7}. In the chromospheric and coronal layers, these fields have splayed out to form the complexes of solar activity (such as active regions and coronal holes), the nature of which varies over the solar cycle.

The magnetic fields observed in the solar atmosphere have their roots deep in the solar interior, where they are presumed to form a predominantly toroidal flux system maintained by dynamo action. Such fields are subject to the vagaries of flux expulsion and topological pumping^{8,9} by convective eddies, and the onset of magnetic buoyancy instabilities⁹. It has been argued^{10–13} that a relatively stable magnetic layer may reside at the base of the convection zone. The field in this layer presumably evolves over the solar cycle, producing, for example, variations in solar luminosity¹¹.

The field strength in the magnetic layer at the base of the convection zone is uncertain, although some reasonable estimates may be made. For the field to resist severe deformation by convective flows, it must be at least as strong as the dynamical pressure, giving a lower limit^{13,14} of $\sim 10^4$ G. At the other extreme, it is unlikely that the field reaches equipartition with the barometric gas pressure; this gives an upper limit of a few times 10^7 G. Also, from a consideration of the partial depletion of ⁷Li inferred from observations, Parker¹⁵ has estimated a field strength of 4×10^5 G just below the base of the convection zone. For the core of the Sun, much stronger fields (10^6 – 10^8 G) have been discussed^{4,16,17}. For the region near the base of the convection zone, though, a range of 10^4 – 10^6 G seems indicated. It is hoped that helioseismology will shed further light on this question.

Whatever the actual magnitude of the field at the base of the convection zone, it presumably varies over the solar magnetic cycle. Such a variation is likely to be reflected in the frequencies of p-modes that penetrate into the magnetic field. Indeed, as we argue here, the decreases in frequencies of low-degree modes from 1980 to 1984 may be due to such a magnetic variation. In this context the recent calculations by Bogdan and Zweibel^{17,18},

Received 6 June; accepted 22 August 1986.

- Gregory, P. C. *et al.* *Nature* **239**, 440–43 (1972).
- Gregory, P. C. & Taylor, A. R. *Astrophys. J.* **248**, 596–605 (1981).
- Taylor, A. R. & Gregory, P. C. *Astr. J.* **88**, 1784–1809 (1983).
- Gregory, P. C. & Taylor, A. R. *Astr. J.* **92**, 371–411 (1986).
- Stepanian, A. A., Vladimirovsky, B. M., Neshpor, Yu. I. & Fomin, V. P. *Astrophys. Space Sci.* **38**, 267–82 (1975).
- Lamb, R. C. & Weekes, T. C. *Astrophys. Lett.* (in the press).
- Bradt, H. V. D. & McClintock, J. E. *A. Rev. Astr. Astrophys.* **21**, 13–66 (1983).
- Stepanian, A. A., Vladimirovsky, B. M. & Fomin, V. P. *Nature phys. Sci.* **239**, 40–41 (1972).
- Chadwick, P. M. *et al.* *Astr. Astrophys.* **151**, L1–3 (1985).
- Gregory, P. C. & Taylor, A. R. *Nature* **272**, 704–706 (1978).
- Gregory, P. C. *et al.* *Astr. J.* **84**, 1030–1036 (1979).
- Taylor, A. R. & Gregory, P. C. *Astrophys. J.* **255**, 210–216 (1982).
- Taylor, A. R. & Gregory, P. C. *Astrophys. J.* **283**, 273–278 (1984).
- Pollock, A. M. *et al.* *Astr. Astrophys.* **94**, 116–120 (1981).

which demonstrate frequency shifts in high-degree p-modes due to the presence of flux-tube (fibril) fields in the photosphere, should be noted.

To examine the influence of a magnetic field on p- and g-modes, we consider the simple circumstances of a plane stratified atmosphere in which a horizontal magnetic field $B(z)\hat{x}$ is embedded. Small-amplitude waves may be described by the equation¹⁹⁻²²

$$\frac{d}{dz} \left\{ \frac{\rho(c_s^2 + v_A^2)(\omega^2 - k^2 c_T^2)}{(\omega^2 - k^2 c_s^2)} \frac{dv_z}{dz} \right\} + \left\{ \rho(\omega^2 - k^2 v_A^2) - \frac{g^2 k^2 \rho}{(\omega^2 - k^2 c_s^2)} - g k^2 \frac{d}{dz} \left(\frac{\rho c_s^2}{\omega^2 - k^2 c_s^2} \right) \right\} v_z = 0 \quad (1)$$

where v_z is the velocity perturbation in the direction of z (aligned vertically, antiparallel to gravity $-g\hat{z}$). The gas density and pressure in the unperturbed atmosphere are $\rho(z)$ and $p(z)$. The sound speed is $c_s = (\gamma p / \rho)^{1/2}$, for constant adiabatic index γ , and the Alfvén speed within the field $B(z)$ is $v_A = B / (4\pi\rho)^{1/2}$; the cusp speed²² is $c_T = c_s v_A / (c_s^2 + v_A^2)^{1/2}$. The wave's frequency and horizontal wavenumber are ω and k , respectively.

The general complexity of equation (1) precludes a specific analytical solution for all but the most special circumstances. However, it is known^{21,22} that equation (1) possesses, in general, both discrete and continuous spectra, associated with fast and slow magnetoacoustic waves respectively. Now, in the absence of a magnetic field, the p-modes of equation (1) form a sturmian sequence, the g-modes an anti-sturmian one²³⁻²⁴. In the presence of a magnetic field, the p-modes of equation (1) form a sturmian sequence, the g-modes an anti-sturmian one^{23,24}. In the presence of a magnetoacoustic mode's continuous spectrum. As a result, no g-modes with discrete frequencies below the cusp frequency $\omega^{\text{cusp}} = k c_T$ can arise. The spectrum of discrete g-modes is thus restricted to frequencies above ω^{cusp} ; below ω^{cusp} this spectrum becomes continuous.

The presence of a magnetic field, then, gives special place to the frequency ω^{cusp} ; an arbitrarily excited disturbance picks out this frequency or, more precisely, the surface-mode frequency, of the field at the base of the convection zone (see ref. 25 and refs therein for a detailed discussion of a related but simpler problem). Of course, the cusp frequency is very low. Taking, for example, $\rho = 0.2 \text{ g cm}^{-3}$ for the gas density at the base of the convection zone²⁶ and $B = 10^5 \text{ G}$, we obtain an Alfvén speed of 0.63 km s^{-1} (compared with a sound speed of $\sim 200 \text{ km s}^{-1}$), which gives a frequency of $\omega^{\text{cusp}} \approx k v_A \approx [l(l+1)]^{1/2} \times 10^{-6} \text{ s}^{-1}$ for mode of degree l . (We have related k to l by $l(l+1) = k^2 R_\odot^2$, where $R_\odot$ is the radius of the Sun.) The cusp frequency for low l is comparable with the solar surface rotation rate.

To see the effect of a uniform magnetic field on the discrete g-modes (having $\omega > \omega^{\text{cusp}}$), consider equation (1) in the low-frequency ($\omega^2 \ll k^2 c_s^2$), weak-field limit ($v_A^2 \ll c_s^2$). Equation (1) may be cast in the form

$$\frac{d^2 \Phi}{dz^2} + \kappa^2(z) \Phi = 0 \quad (2)$$

where $\Phi = \rho^{1/2}(\omega^2 - k^2 v_A^2)^{1/2} v_z$, and

$$\kappa^2(z) \approx -k^2 + \frac{k^2 \omega_g^2}{(\omega^2 - k^2 v_A^2)} + \frac{\omega^2 k^2 v_A^2}{4H^2(\omega^2 - k^2 v_A^2)^2} \quad (3)$$

The buoyancy frequency ω_g is given by

$$\omega_g^2 = \frac{g}{H} - \frac{g^2}{c_s^2}, \quad H^{-1} = -\frac{1}{\rho} \frac{d\rho}{dz} \quad (4)$$

in a stratified atmosphere with density scale-height H .

Comparing this with the zero-field case ($\kappa^2 \approx k^2 \omega_g^2 / \omega^2$) we note the additional wavenumber dependence arising through the denominator $\omega^2 - k^2 v_A^2$. The g-mode cavity ($\kappa^2 = 0$) is seen

to be dependent on ω_g , ω , k and v_A , in contrast to the field-free case when the cavity is determined by the buoyancy frequency alone. Evidently, the presence of a magnetic field modifies the property²⁷⁻³¹ of uniform spacing with order n of the g-mode periods ($2\pi/\omega_n$). In a weak field this effect is not large, except for frequencies ω near to $k v_A$. Assuming an effective k given by $k^2 = l(l+1)/R_\odot^2$, the frequency $k v_A$ is very small for low-degree modes; $k v_A / 2\pi \approx 2 \mu\text{Hz}$ in a field of 10^6 G . However, the effective k may be much larger ($k^2 = l(l+1)/r^2$ for radius r) than its surface value (where $r = R_\odot$), in which case the influence of a magnetic field on g-modes will be more pronounced. Overall, we may expect the field to cause an increase in ω over its field-free value, the increase being dependent on wavenumber. To fully assess this effect, however, requires a model of the g-modes in a magnetic sphere, which is unavailable to us.

We turn now to a consideration of the p-modes. Rather than use equation (1), it is convenient to follow the example of the non-magnetic case and work in terms of the variable $\text{div } \mathbf{v}$. As in the non-magnetic case, $\text{div } \mathbf{v}$ satisfies a second-order differential equation in z , the coefficients of which are unfortunately algebraically complicated. However, for the circumstances of interest here, where $\omega \gg k v_A$, we may expand the coefficients in our equation in powers of $k^2 v_A^2 / \omega^2$. The result is an equation of the form (2) where $\Phi \approx \rho^{1/2}(c_s^2 + v_A^2) \text{div } \mathbf{v}$, and

$$\kappa^2 = -k^2 + \frac{\omega^2}{c_s^2 + v_A^2} - \frac{\omega_c^2}{c_s^2} + \frac{k^2 c_s^2 v_A^2}{(c_s^2 + v_A^2)^2} + \frac{k^2 \omega_g^2}{\omega^2} \left(\frac{c_s^2}{c_s^2 + v_A^2} \right) + A \quad (5)$$

Here ω_c is the Lamb cutoff frequency³² given by $\omega_c^2 = c_s^2(1 - 2dH/dz)/4H^2$. The magnetic field strength B is assumed uniform.

The coefficient A in equation (5) is of order $k^2 v_A^2 / (\omega^2 H^2)$ and is generally negligible; it is identically zero for k , v_A or g equal to zero. When $v_A = 0$ equation (5) recovers Lamb's³² result; see refs 27-31 for a discussion of the helioseismological implications of equations (2) and (5) in the zero-field case. Note that expression (5) is not obtainable from its field-free counterpart simply by replacing c_s^2 with $(c_s^2 + v_A^2)$. A detailed derivation of equations (2) and (5) will be given elsewhere (work in preparation).

As with the non-magnetic case, equation (2) is amenable to a WKB (Wentzel-Kramers-Brillouin) analysis, yielding²⁷⁻³¹

$$\int_{\text{cavity}} \kappa dz \approx (n + \varepsilon)\pi \quad (6)$$

for phase ε ; the integration in equation (6) is over the cavity of the mode, defined by the zeros of $\kappa(z)$.

For the high-frequency p-modes we follow the non-magnetic case and neglect terms in ω_g^2 ($\omega_g / 2\pi$ is $\leq 0.5 \text{ mHz}$) and ω_c^2 (the contribution from which may be absorbed into the phase²⁸, ε). Also, for comparison with the field-free results, it is convenient to write our expressions in terms of radius r , replacing k^2 by $l(l+1)/r^2$. Then equation (6) becomes

$$\int_{\text{cavity}} \left[\frac{\omega^2}{c_s^2} - \frac{l(l+1)}{r^2} \right]^{1/2} \left(1 + \frac{v_A^2}{c_s^2} \right)^{-1/2} dr \approx \pi(n + \varepsilon) \quad (7)$$

It is clear from equation (7) that the effect of a non-zero field is to increase the frequency ω above its value in the absence of a field.

For p-modes of low degree penetrating deep into the magnetic field, the cavity is essentially the whole of the solar interior. It is sufficient for our purposes to represent v_A^2/c_s^2 by its mean value, β^{-1} , over the integration range ($r = 0$ to $R_\odot$). Then, in terms of cyclic frequency $\nu_B (\equiv \omega/2\pi)$, equation (7) with $\beta^{-1} \ll 1$ yields

$$\nu_B = (n + \frac{1}{2}l + \varepsilon) \Delta\nu_0 \left(1 + \frac{1}{2\beta} \right) \quad (8)$$

where $\Delta\nu_0 = (2 \int_0^{R_\odot} dr / c_s)^{-1}$.

Let ν_0 be the oscillation frequency determined by equation (8) in the absence of a magnetic field (that is, when $\beta^{-1} = 0$). Then, in the presence of a magnetic field, the frequency ν_B is increased by an amount

$$\nu_B - \nu_0 = \frac{1}{2\beta} \nu_0 \quad (9)$$

Thus, the frequency increase due to the presence of a magnetic field is proportional to the square of the magnetic field strength. A similar conclusion has been reached by Bogdan and Zweibel^{17,18} for fibril magnetic fields (see also ref. 33). This frequency increase is, of course, very small; for example, $\nu_B - \nu_0 = 5 \times 10^{-6} \nu_0$ for a field strength $B = 10^5$ G.

Consider, however, the change in frequency due to a change in magnetic field strength over the solar cycle. If we define $\Delta\nu_B \equiv \nu_B - \nu_{B'}$, the difference in frequencies in magnetic fields B and B' , then

$$B^2 - B'^2 = 8\pi\rho c_s^2 \frac{\Delta\nu_B}{\nu_0} \quad (10)$$

With $\rho = 0.2 \text{ g cm}^{-3}$, $c_s = 2 \times 10^7 \text{ cm s}^{-1}$ and $\nu_0 = 3.1 \text{ mHz}$ we obtain

$$B^2 - B'^2 = 6.5 \times 10^{17} \Delta\nu_B \text{ G}^2 \quad (11)$$

Woodard and Noyes⁵ have recently reported a decrease in low-degree p-mode frequencies, from 1980 to 1984, of $0.42 \mu\text{Hz}$. Associating this frequency decrease with a decline in field strength from a value B in 1980 (close to solar maximum) to a value B' in 1984 (approaching solar minimum), we see from equation (11) that

$$B^2 - B'^2 = 2.7 \times 10^{11} \text{ G}^2 \quad (12)$$

Thus, for the magnetic field to be responsible for the changes reported in ref. 5, it is necessary that the peak field strength at solar maximum be $\geq 5 \times 10^5$ G. The actual variation in field strength necessary to explain the observed frequency shift depends, of course, on the magnitude of the peak field. For example, with a peak field of 7×10^5 G a decline to 5×10^5 G reproduces the frequency decrease; for a peak field of 10^6 G, a decline to 8.6×10^5 G is necessary.

The above estimates of field strengths are too low if the varying field at the base of the convection zone is in a fibril state. Returning to equation (7) and supposing that v_A is non-zero only within a layer of thickness d , we see that the right-hand side of equation (9) is multiplied by a filling factor $f \approx 2d\Delta\nu_0/c_s$ (see also refs 17, 18). The right-hand sides of equations (10) and (11), in turn, are multiplied by f^{-1} . An estimate of $d = 10^4$ – 10^5 km gives $f = 10^{-2}$ – 10^{-1} . With $f = 10^{-1}$, equation (11) gives a peak field strength of $\geq 1.6 \times 10^6$ G; with $f = 10^{-2}$, the peak field is $\geq 5 \times 10^6$ G.

All of these estimates, based on the WKB evaluation of the frequencies and an approximate treatment of the magnetic corrections to the integrals, are admittedly crude. Nonetheless, they are borne out by a detailed numerical analysis (work in preparation) of the frequency shifts for a polytropic atmosphere with a uniform magnetic field (of arbitrary strength) at the base of the convection zone; peak field strengths of $\sim 2 \times 10^6$ G are required to reproduce the observed frequency decrease.

We thank Professor Donald Gurnett and his colleagues for hospitality at the University of Iowa. This work was supported in part by NASA through grants NGL-16-001-043 & NSG-7411. W.R.C. acknowledges financial support from the Department of Education for Northern Ireland.

Note added in proof: New ground-based data presented by Elsworth *et al.*³⁴ cast some doubt on the trend in the p-mode frequencies reported in ref. 5, and emphasize the need for continued acquisition of data over a complete solar cycle.

3. Christensen-Dalsgaard, J., Duvall, T. L. Jr, Gough, D. O., Harvey, J. W. & Rhodes, E. J. *Jr Nature* **315**, 378–382 (1985).
4. Dziembowski, W. & Goode, P. R. *Nature* **305**, 39–42 (1983).
5. Woodard, M. F. & Noyes, R. W. *Nature* **318**, 449–450 (1985).
6. Zwaan, C. in *The Sun as a Star* (ed. Jordan, S.) Ch. 6 (NASA SP-450, 1981).
7. Spruit, H. C. & Roberts, B. *Nature* **304**, 401–406 (1983).
8. Moffatt, H. K. *Magnetic Field Generation in Electrically Conducting Fluids* (Cambridge University Press, 1978).
9. Parker, E. N. *Cosmical Magnetic Fields* (Clarendon, Oxford, 1979).
10. Parker, E. N. *Astrophys. J.* **198**, 205–209 (1975).
11. Spiegel, E. A. & Weiss, N. O. *Nature* **287**, 616–617 (1980).
12. Galloway, D. J. & Weiss, N. O. *Astrophys. J.* **243**, 945–953 (1981).
13. Van Ballegoijen, A. *Astr. Astrophys.* **106**, 43–52; **113**, 99–112 (1982).
14. Schüssler, M. in *Solar and Stellar Magnetic Fields: Origins and Coronal Effects* (ed. Stenflo, J. O.) 213–236 (Reidel, Dordrecht, 1983).
15. Parker, E. N. *Astrophys. J.* **286**, 666–676; 677–679 (1984).
16. Dicke, R. H. *Astrophys. J.* **228**, 898–902 (1979).
17. Bogdan, T. J. & Zweibel, E. G. *Astrophys. J.* **298**, 867–875 (1985).
18. Zweibel, E. G. & Bogdan, T. J. *Astrophys. J.* (in the press).
19. Nye, A. & Thomas, J. H. *Astrophys. J.* **204**, 573–581 (1976).
20. Adam, J. A. *Sol. Phys.* **52**, 293–307 (1977).
21. Goedbloed, J. P. *Physica* **12D**, 107–132 (1984).
22. Roberts, B. in *Solar System Magnetic Fields* (ed. Priest, E. R.) Ch 3 (Reidel, Dordrecht, 1985).
23. Spiegel, E. A. & Unno, W. *Publ. astr. Soc. Jap.* **14**, 28–32 (1962).
24. Christensen-Dalsgaard, J. *Mon. Not. R. astr. Soc.* **190**, 765–791 (1980).
25. Lee, M. A. & Roberts, B. *Astrophys. J.* **301**, 430–439 (1986).
26. Spruit, H. C. *Sol. Phys.* **34**, 277–290 (1974).
27. Deubner, F.-L. & Gough, D. A. *Rev. Astr. Astrophys.* **22**, 593–619 (1984).
28. Christensen-Dalsgaard, J. in *Theoretical Problems in Stellar Stability and Oscillations* (eds Gabriel, M., Noels, A.) 155–207 (Institut d'Astrophysique, Liege, 1984).
29. Christensen-Dalsgaard, J., Gough, D. & Toomre, J. *Science* **229**, 923–931 (1985).
30. Leibacher, J. W., Noyes, R. W., Toomre, J. & Ulrich, R. K. *Scient. Am.* **253**, 48–57 (1985).
31. Gough, D. *Sol. Phys.* **100**, 65–99 (1985).
32. Lamb, H. *Hydrodynamics* (Cambridge University Press, 1932).
33. Carroll, B. W. *et al. Astrophys. J.* **305**, 767–783 (1986).
34. Elsworth, Y. P. *et al.*, in *Helioseismology*, IAU Colloq. 123 (Aarhus, Denmark, July 1986, preprint).

Radio detection of uranian lightning by Voyager 2

P. Zarka & B. M. Pedersen

Groupe de Radioastronomie Décimétrique, UA CNRS 324, Observatoire de Paris, Section de Meudon, F-92195 Meudon Principal Cedex, France

Within distances of $\sim 600,000$ km from Uranus, the Planetary Radio Astronomy (PRA) experiment aboard the Voyager 2 spacecraft detected impulsive (100–300 ms) bursts of broad-band (≤ 900 kHz to ~ 40 MHz) radio emission. This emission is very different from the uranian magnetospheric radio component, also discovered during the recent Voyager–Uranus encounter¹. During the two Saturn encounters, in 1980 and 1981, the PRA experiments detected similar radio emissions^{2,3}, which have been interpreted as radio emissions associated with lightning activity in the atmosphere of the planet^{4,5}. By analogy with the Saturn electrostatic discharges (SED), we have adopted the term uranian electrostatic discharges (UED) for these new radio bursts. We have determined the physical characteristics of UED, and in the context of their interpretation as lightning-associated radio bursts, we derive here ionospheric peak electron densities over the uranian day- and nightside hemispheres.

Voyager 2's closest approach to Uranus ($\sim 80,000$ km above the cloud tops) took place on 24 January 1986 at 17:59. Figure 1 displays an important cluster of UED, as they appear on a dynamic spectrum recorded by the PRA experiment⁶ ~ 1 h after closest approach. This dynamic spectrum is generated by successive 6-s scans of the 198 channels of the PRA receiver, grouped in two frequency bands: a high-frequency (HF) band comprising 128 channels with a 200-kHz bandwidth uniformly spaced from 40.5 to 1.2 MHz, and a low-frequency (LF) band comprising 70 channels with a 1-kHz bandwidth uniformly spaced between 1,326 and 1.2 kHz. A measurement of the intensity received in one channel takes 30 ms, and is made alternatively in left- and right-hand circular polarization. In Fig. 1, increasing darkness is proportional to increasing intensity. As for Saturn, any sort of spacecraft interference or nearby discharging as the source of these bursts can be ruled out, and similarly to SED, the UED can be interpreted as

Received 27 May; accepted 21 July 1986.

1. Duvall, T. L. Jr *Nature* **300**, 242–243 (1982).
2. Duvall, T. L. Jr & Harvey, J. W. *Nature* **302**, 24–27 (1983).

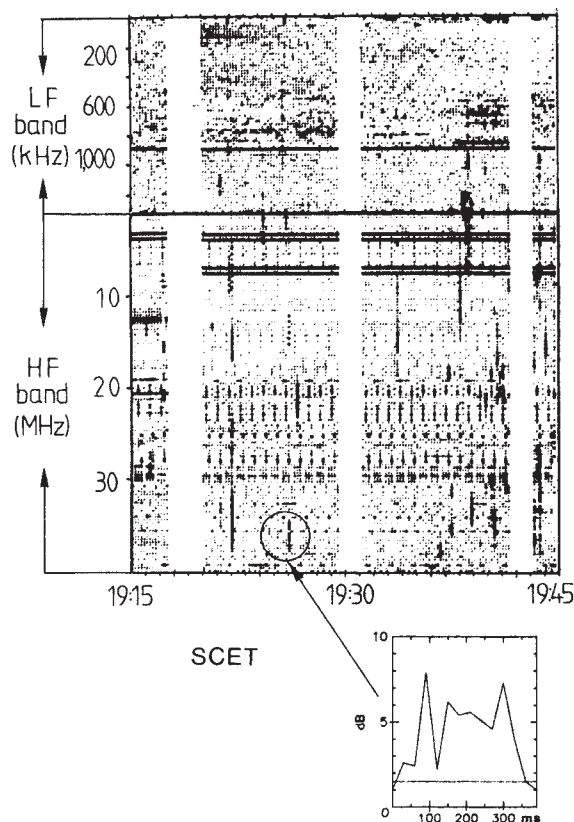


Fig. 1 PRA dynamic spectrum in high- and low-frequency bands, recorded ~ 1 h after closest approach. SCET, spacecraft event time (Universal Time at Voyager 2). Horizontal lines and regular vertical streaks are spacecraft interference. Uranus magnetospheric emission appears in the LF band below ~ 800 kHz. Vertical white areas are data gaps. As UED are broad-band phenomena, with durations short (100–300 ms) compared with the sweeping rate of the PRA receiver (6 s), they appear as vertical streaks randomly distributed over the spectrum. The profile of a typical burst (circled) is also shown, as a function of the time in the scan at which it occurs. Polarized UED events appear as dotted lines on the dynamic spectrum, for example, at $\sim 19:26$.

broad-band bursts detected only in those PRA channels which are sampling emission during the short-lived interval when they occur. Their signatures thus appear randomly distributed over the PRA frequency range. The profile of an individual UED event is also given in Fig. 1, showing rapid fluctuations over 5–10 dB of the received flux.

Figure 2a displays the ~ 140 UED events detected during the Voyager 2–Uranus flyby. To identify these events, we have used an improved version of the computer algorithm written for SED recognition⁷, which is much more sensitive than the visual counting of UED on the dynamic spectra. Only bursts covering three channels or more (and therefore lasting for at least 90 ms) have been taken into account, and several PRA channels which were particularly noisy due to spacecraft interference have been eliminated. Therefore, UED are absent from the frequency intervals indicated on Fig. 2b. In the LF band, the highly fluctuating background and especially the presence of uranian kilometric radiation¹ (UKR) up to 800 kHz prevents an automatic detection of UED. Thus, although weak low-frequency events were detected on day 25, we retained only the close-encounter events, whose existence has been confirmed visually from dynamic spectra such as Fig. 1. Considering these restrictions, together with the variable UED low-frequency cutoff discussed below, and the frequency response curve of the PRA receiver and antennas⁸, the rather uniform burst distribution over the PRA frequency range suggests a broad and smooth spectrum for UED over the range ≤ 900 kHz–40 MHz. The occurrence diagram deduced from Fig. 2a is shown in Fig. 2c. Although UED tend to group into episodes, no recurrence period is evident. Note that in Fig. 2a, UED at frequencies above ~ 25 MHz, or below ~ 7 MHz, were detected almost solely during the episode near closest approach, which also contains more than half of all detected events.

A few UED events seem to be polarized, as they appear as dotted lines on the dynamic spectra (that is, they are detected only by succeeding channels of the same polarization sense); an example is the event near 19:26 in Fig. 1. These events, which appear with either polarization sense, are not included in Fig. 2 because our identification algorithm retains only those UED bursts covering at least three consecutive channels. However, the retained events cover an approximately equal number of left- and right-hand-polarized PRA channels. This suggests that, as was the case for SED, the UED are intrinsically unpolarized,

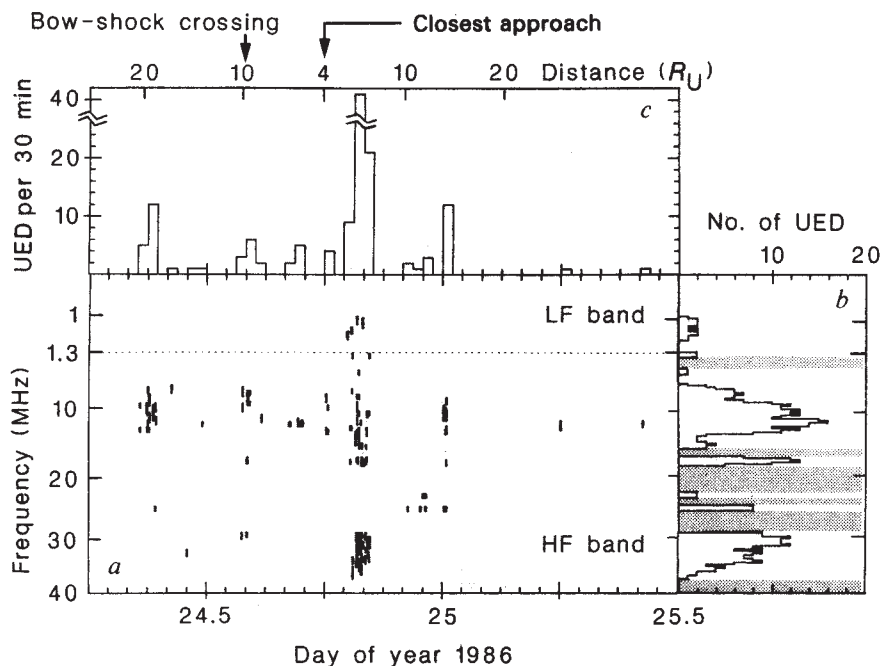


Fig. 2 a, Distribution in the time-frequency plane of all the UED detected during the Voyager 2–Uranus encounter. b, Number of events observed per PRA channel (obtained through time summation of Fig. 2a). Grey-shaded areas correspond to the channels that have been eliminated due to strong interference, hence the absence of UED in the corresponding frequency intervals on Fig. 2a. c, UED occurrence diagram. The number of events in 30-min bins is plotted versus time and distance to Uranus.

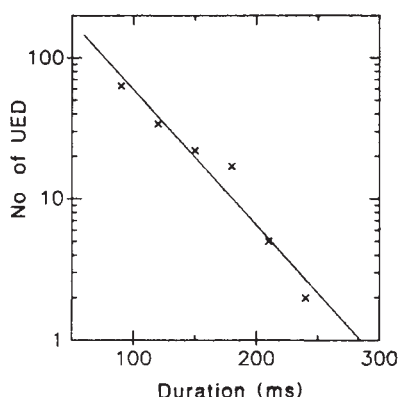


Fig. 3 Distribution of UED durations, and best fit by an exponential law with an e-folding time of ~ 100 ms.

but may occasionally exhibit an apparent polarization due to the geometry of the PRA antennas relative to the UED source location, and possibly due to propagation phenomena.

Figure 3 displays the distribution of durations of individual UED events, which are directly proportional to their apparent lengths on the dynamic spectra. As for SED⁷, this distribution is very well described (with a $>99\%$ degree of confidence) by an exponential law of the form $N(D) = N_0 \exp(-D/D_0)$, with an e-folding time $D_0 \approx 100$ ms, and an average duration of ~ 120 ms. These characteristic durations are about twice as long as for SED. A few bursts, not included in Fig. 3, last for more than 300 ms (up to 500 ms).

The UED spectrum is shown in Fig. 4. It has been calculated from the averaged intensities of the detected bursts in each clean PRA channel, normalized to a distance of 1 AU, and including the correction for the response curve of the PRA receiver^{7,8}. The broad-band UED spectrum decreases smoothly from low to high frequencies (approximately as f^{-2} , as for radio emissions associated with terrestrial lightning⁹). The average intensity is $6 \times 10^{-24} \text{ W m}^{-2} \text{ Hz}^{-1}$ in the HF band, and $2 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$ in the LF band. For a source close to Uranus with an isotropic radiation pattern, these intensities correspond to spectral powers of respectively 2 and 60 W Hz^{-1} . Thus, the total instantaneous power, integrated over the whole UED spectrum, is $\sim 10^8 \text{ W}$. By comparison, the SED spectrum was found to be smooth over the whole PRA HF band, the intensity decrease probably occurring at even higher frequencies, outside the frequency range of the receiver. The estimated SED instantaneous power is ≥ 10

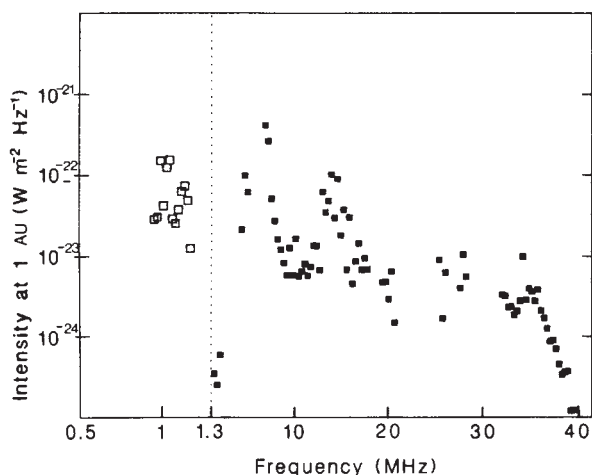


Fig. 4 UED spectrum normalized to 1 AU and corrected for instrumental effects. The frequency scales are linear, but different for LF band ($\square$) and HF band ($\blacksquare$).

times higher than for UED. Note that the detection threshold in the PRA HF band ($\sim 6 \times 10^{-19} \text{ W m}^{-2} \text{ Hz}^{-1}$) should make average UED bursts detectable from distances $\leq 20 R_u$ ($1 R_u = 1$ uranian radius = 25,600 km), which is the case in Fig. 2a. This again confirms the uranian origin of these bursts. Towards high frequencies, this detection threshold, together with the decrease of UED intensities and of the response of the PRA receiving system⁸, explains the lack of detection of UED above 25 MHz, except for a short period very near closest approach. At low frequencies before closest approach, however, due to the high burst intensities and the high gain of the PRA receiving system, the non-detection of UED should be attributed to a cutoff effect.

In spite of the small number (~ 140) of UED events detected by Voyager 2 during the unique Uranus encounter, compared with the 23,000 SED observed during the two consecutive Saturn encounters⁷, their physical properties, as well as their similarity to SED (and terrestrial lightning) allow us to conclude that UED are natural radio emissions from Uranus. We interpret them as lightning-associated radio bursts originating from storms in the uranian atmosphere. However, there is a controversy over the origin of SED. An alternative interpretation is the ring source hypothesis, which locates the source of SED at 1.8 Saturn radii in the middle of Saturn's B-ring^{2,3,10}. However, this explanation does not account for the observed SED characteristics¹¹. Thus, in the context of an atmospheric source, the variations of UED's LF limit of appearance from before to during encounter is simply explained in terms of ionospheric cutoff. The value of the cutoff gives an upper limit for the peak electron plasma frequency in the sub-spacecraft ionosphere, below which no radio wave can escape from the atmosphere towards the overflying spacecraft. Before encounter, Voyager 2 was above the dayside hemisphere of the planet. The UED low-frequency cutoff at ~ 7 MHz thus implies a peak ionospheric electron density $N_e \approx 6 \times 10^5 \text{ cm}^{-3}$ over this hemisphere. In the same way, the low-frequency cutoff at ≤ 900 kHz during closest approach corresponds to a peak electron density $N_e \leq 10^4 \text{ cm}^{-3}$ in the nightside ionosphere. These results are consistent with the preliminary conclusions of the radio science team¹², who estimate from Voyager 2 immersion data an ionospheric electron density, near the planet's limb, of several thousand electrons per cm^3 .

Compared with the day/night ratio of peak ionospheric electron densities measured on Saturn and the Earth (up to several hundred^{11,13}), the difference by a factor of 60 in the case of Uranus could seem to be astonishingly low, especially for a planet having the same hemisphere directed towards (or away from) the Sun for years. A possible explanation might be that electrons generated by sunlight in the high-density dayside ionosphere are constantly drifting to the nightside hemisphere along magnetic flux tubes having their footprints in either hemisphere. However, due to the geometry of the tilted uranian magnetic dipole field¹⁴, such a transport process should directly affect only low-latitude regions of the nightside ionosphere, ($\leq 30^\circ$). The difference in electron densities between sunlit and dark hemispheres, as well as between low and high nightside latitudes, could have important consequences for the dynamics of the uranian magnetosphere.

With an instantaneous power of $\sim 10^8 \text{ W}$ and a mean duration of 120 ms, the typical energy radiated at radio wavelengths per uranian lightning flash is $(1-2) \times 10^7 \text{ J}$. In terms of the energy emitted per flash, uranian lightning are thus comparable to jovian lightning, much more powerful than the terrestrial ones, but ~ 10 times less so than SED¹⁵. The scarcity of UED detected during the Uranus flyby may partly be explained by a low, but non-negligible, level for the convective motions which are necessary for lightning generation. However, this remains to be confirmed by the recent Voyager 2 observations. Another explanation comes from the composition of the uranian atmosphere, whose measured main cloud constituents are hydrogen, helium, methane and acetylene^{17,18}, all unfavourable candidates

to support electrification, due to their symmetrical structure and therefore low polarizability (or dielectric constant)¹⁶. However, a detailed discussion of the implications of this discovery for our understanding of the planet's atmosphere and magnetosphere would be speculative, as it is based on data from a unique Voyager flyby.

We thank M. L. Kaiser for useful remarks and the Centre National d'Etudes Spatiales for support.

Received 19 May; accepted 1 August 1986.

1. Warwick, J. W. *et al. Science* **233**, 102-106 (1986).
2. Warwick, J. W. *et al. Science* **212**, 239-243 (1981).
3. Warwick, J. W. *et al. Science* **215**, 582-587 (1982).
4. Burns, J. A., Showalter, M. R., Cuzzi, J. N. & Durisen, R. H. *Icarus* **54**, 280-295 (1983).
5. Kaiser, M. L., Connerney, J. E. P. & Desch, M. D. *Nature* **303**, 50-53 (1983).
6. Warwick, J. W., Pearce, J. B., Peltzer, R. G. & Riddle, A. C. *Space Sci. Rev.* **21**, 309-327 (1977).
7. Zarka, P. & Pedersen, B. M. *J. geophys. Res.* **88**, 9007-9018 (1983).
8. Ortega-Molina, A. & Daigne, G. *Astr. Astrophys.* **130**, 301-310 (1984).
9. Volland, H. in *Handbook of Atmospherics*, Vol. 1 (ed. Volland, H.) 179-250 (CRC Press, Boca Raton, Florida, 1982).
10. Evans, D. R., Romig, J. H. & Warwick, J. W. *Icarus* **54**, 267-279 (1983).
11. Zarka, P. thesis, Univ. Paris VI (1984); in *Planetary Radio Emissions* (eds Rucker, H. O. & Bauer, S. J.) 237-270 (Verlag der Österreichischen Akademie der Wissenschaften, 1985).
12. Tyler, G. L. *et al. Science* **233**, 79-84 (1986).
13. Davies, K. *Ionospheric Radio Propagation* (NBS Monogr. (US) **80**, Washington DC, 1965).
14. Ness, N. F. *et al. Science* **233**, 85-89 (1986).
15. Zarka, P. *Astr. Astrophys.* **146**, L15-L18 (1985).
16. Rinnert, K. *J. geophys. Res.* **90**, 6225-6237 (1985).
17. Smith, B. A. *et al. Science* **233**, 43-64 (1986).
18. Broadfoot, A. L. *et al. Science* **233**, 74-79 (1986).

Quantum jumps and atomic cryptograms

P. L. Knight*, R. Loudon† & D. T. Pegg*‡

* Optics Section, Blackett Laboratory, Prince Consort Road, London SW7 2BZ, UK

† Physics Department, Essex University, Colchester CO4 3SQ1, UK

Can a single atom coherently excited by on-resonance laser radiation generate a random telegraph signal with frequent periods of darkness in the fluorescence radiation, reflecting the quantum jumps of the atomic transitions? Erber and Putterman¹ have proposed that such a random telegraph, predicted earlier² for incoherent excitation, persists when the excitation is coherent and resonant with the unperturbed atomic transition frequencies. They argue that such an atomic telegraph is cryptographically equivalent to an infinite computer in its ability to generate random numbers, and would provide new tests of basic quantum theory. Here we argue that any periods of darkness in the light emitted by a coherently driven atom excited on resonance must be exceedingly rare.

Figure 1 shows the three-level atom in the 'vee' configuration discussed in refs 1 and 2, with levels S, O and W simultaneously irradiated. The excitation from the ground state O to the excited state S is considered to be very strong, whereas that to the excited state W is considered to be weak. When the radiation is incoherent there are periods of darkness in the strong transition fluorescence which are of the order of magnitude of the bright periods, provided the weak transition excitation rate is greater than the weak decay rate A_2 , giving a significant reduction in the overall strong transition fluorescence. It is these that constitute the random telegraph. This has been demonstrated both by a coarse-grained time-averaging approach² and also in terms of the individual photon events³. In the latter theory, the periods of darkness are manifestations of the cross-antibunching effect when the strong-transition photons are followed by weak-transition photons. Mathematically, this is exhibited by the long rise-time of the correlation function for detection of a weak-transition photon, denoted by subscript 2, a time τ after detection of a strong-transition photon, denoted by subscript 1. The corre-

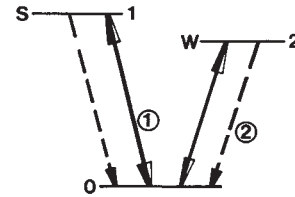


Fig. 1 Energy-level scheme for the three-level atomic 'vee' configuration discussed in the text. The transition between S and O is strongly driven; that between W and O is weakly driven. The lines with double arrows represent the driving fields and the dashed lines the spontaneous emission transitions. The strong coupled state is 1 and the weak coupled state is 2. ①, Strong-transition photon; ②, weak-transition photon.

sponding normalized correlation function, or degree of second-order coherence $g_{21}^{(2)}(\tau)$, is calculated by a standard Einstein rate-equation analysis⁴ for driving of the three-level atom by two incoherent fields. We assume that the driving rate $B_1 W_1$ of the strong transition (where W_1 is the spectral energy density of the incident light at the strong-transition frequency) is much larger than its spontaneous decay rate A_1 and the corresponding rates $B_2 W_2$ and A_2 for the weak transition. Then

$$g_{21}^{(2)}(\tau) = 1 - \frac{(3B_2 W_2 + A_2)}{2B_2 W_2} \exp(\lambda_+ \tau) + \frac{(B_2 W_2 + A_2)}{2B_2 W_2} \exp(\lambda_- \tau) \quad (1)$$

where the fast and slow decay rates are

$$\lambda_+ \approx -[2B_1 W_1 + A_1 + \frac{1}{2}B_2 W_2] \quad (2)$$

and

$$\lambda_- \approx -[\frac{3}{2}B_2 W_2 + A_2] \quad (3)$$

This and the other correlation functions $g_{ij}^{(2)}(\tau)$, where $i, j = 1, 2$, are shown in Fig. 2.

A period of darkness in the strong-transition fluorescence heralds the emission of a weak-transition photon. The introduction of the weak-transition irradiation has a marked effect on the strong-transition fluorescence, as can be seen by calculating the steady-state occupation probabilities of the states. With the level notation shown in Fig. 1, these are

$$\rho_{00} = \rho_{11} = \frac{(A_2 + B_2 W_2)}{(2A_2 + 3B_2 W_2)} \quad (4)$$

$$\rho_{22} = \frac{B_2 W_2}{(2A_2 + 3B_2 W_2)} \quad (5)$$

where $B_1 W_1$ has been assumed to be very much larger than the other transition rates. In the absence of any radiation at the weak-transition frequency ($W_2 = 0$), the O-S transition is fully saturated, with ground- and strong-excited-state occupation probabilities equal to $\frac{1}{2}$. However, these probabilities decrease when W_2 is non-zero, and they drop to $\frac{1}{3}$ when $B_2 W_2 \gg A_2$. This

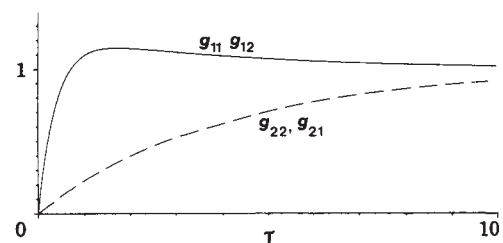


Fig. 2 Degrees of second-order coherence, $g_{ij}^{(2)}(\tau)$, with $i, j = 1, 2$, plotted as a function of τ with parameter values $B_1 W_1 = 1$, $B_2 W_2 = 0.1$, $A_1 = 0.5$, $A_2 = 0.1$.

‡ Permanent Address: School of Science, Griffith University, Nathan, Brisbane 4111, Australia.

occurs even in the limit when the O-S transition rate approaches infinity, with a corresponding reduction in the long-time average fluorescence intensity. Javanainen⁵ has calculated the appropriate correlation functions for a very simple multi-level atomic decay and has compared his results with those of a classical two-state telegraph.

When on-resonance coherent irradiation by two lasers is used instead of incoherent irradiation, we find that the effect described above (and claimed in ref. 1 to still occur in the coherent case) is radically changed. Again this is most easily seen by calculating the long-time average occupation probability of state S. For coherent excitation, the three-level 'vee' atomic system is described by the Bloch equations⁴

$$\dot{\rho}_{11} = i\alpha_1(\rho_{10} - \rho_{01}) - A_1\rho_{11} \quad (6)$$

$$\dot{\rho}_{22} = i\alpha_2(\rho_{20} - \rho_{02}) - A_2\rho_{22} \quad (7)$$

$$\dot{\rho}_{12} = i\alpha_2\rho_{10} - i\alpha_1\rho_{02} - \frac{1}{2}(A_1 + A_2)\rho_{12} \quad (8)$$

$$\dot{\rho}_{10} = i\alpha_1(\rho_{11} - \rho_{00}) + i\alpha_2\rho_{12} - \frac{1}{2}A_1\rho_{10} \quad (9)$$

$$\dot{\rho}_{20} = i\alpha_2(\rho_{22} - \rho_{00}) + i\alpha_1\rho_{21} - \frac{1}{2}A_2\rho_{20} \quad (10)$$

where the Rabi frequencies α_1 and α_2 replace the incoherent energy densities used earlier. The complete steady-state solutions of these equations are complicated, but in the limit in which A_2 is much less than any other rate or frequency, we find in steady state as time $t \rightarrow \infty$

$$\rho_{00} \approx \frac{1}{2} \quad (11a)$$

$$\rho_{11} \approx \frac{1}{2} \frac{\alpha_1^2}{\alpha_1^2 + \alpha_2^2 + \frac{1}{4}A_1^2} \quad (11b)$$

$$\rho_{22} \approx \frac{1}{2} (\alpha_2^2 + \frac{1}{4}A_1^2) / [\alpha_1^2 + \alpha_2^2 + \frac{1}{4}A_1^2] \quad (11c)$$

When α_1 is much larger than all the other transition rates, these equations further reduce to

$$\rho_{00} \approx \rho_{11} \approx \frac{1}{2}, \quad \rho_{22} \approx 0 \quad (12)$$

These results for coherent excitation are to be compared with the results (4) and (5) for incoherent excitation³ (see also refs 6, 7 for further work on rate-equation incoherent excitation, and ref. 8 for a dressed-atom treatment of coherent excitation). The coherently and resonantly excited level populations show none of the dependence on the strength of the radiation at the weak-transition frequency found in the incoherent case. The occupation probability of state S is unaffected by the introduction of weak-transition radiation, corresponding to a non-zero α_2 , always assuming that the O-S Rabi frequency is very large. Thus, any periods of darkness must be exceedingly rare. The resonance frequencies in such a situation are split into intensity-dependent Autler-Townes doublets; Dalibard and Cohen-Tannoudji⁸ have discussed quantum jumps when the weak laser is tuned to these new frequencies. Indeed, the Autler-Townes effect is an alternative way of looking at the problem we have addressed; the Autler-Townes splitting caused by the strong-transition laser will, if $\alpha_1 \gg \alpha_2$, shift the weak-transition resonance away from the unperturbed atomic frequency by an amount greater than the power width of the weak transition. This reduces the effect of the weak-transition laser, which is set at the unperturbed atomic frequency.

The large quantitative difference between the result predicted by Erber and Putterman¹ for their proposed on-resonance coherent excitation experiment and those predicted by conventional quantum mechanics supports to some extent the claim by Erber and Putterman that their experiment will provide a test of the fundamental principles of quantum theory, but the principle at stake seems to be the well-tried one of coherence in superposition states.

We thank M. A. Lauder for computer graphics, C. Cohen-Tannoudji, J. Dalibard, J. Javanainen and S. Stenholm for discussions and the SERC and the British Council for their support.

Received 30 January 1986; accepted 4 July 1986.

1. Erber, T. & Putterman, S. *Nature* **318**, 41-42 (1985).
2. Cook, R. J. & Kimble, H. J. *Phys. Rev. Lett.* **54**, 1023-1026 (1985).
3. Pegg, D. T., Loudon, R. & Knight, P. L. *Phys. Rev. A* **33**, 4085-4091 (1986).
4. Loudon, R. *The Quantum Theory of Light*, 2nd edn (Clarendon, Oxford, 1983).
5. Javanainen, J. *Phys. Rev. A* **33**, 2121-2124 (1986).
6. Arecchi, F. T., Schenzle, A., DeVoe, R. G., Jungmann, K. & Brewer, R. G. *Phys. Rev. A* **33**, 2124-2126 (1986).
7. Schenzle, A., DeVoe, R. G. & Brewer, R. G. *Phys. Rev. A* **33**, 2127-2131 (1986).
8. Dalibard, J. & Cohen-Tannoudji, C. *Europhys. Lett.* **1**, 441-448 (1986).

Do climatic attractors exist?

Peter Grassberger

Physics Department, University of Wuppertal, D-56 Wuppertal, FRG

Recently, it has been claimed¹ that the worldwide climate over the past million years follows a low-dimensional strange attractor. Contrary to that claim, I report here that there is no sign of such an attractor. This holds both for the worldwide climate of the past 1-2 Myr (averaged over periods of ~5,000 yr) and for a local climate of the past ~7,000 yr with yearly averages. In this context, I shall also discuss the general problem of dimension estimates from sparse data.

In view of the pronounced variability of climates, expressed most markedly in the succession of ice ages of the past million years, there have been numerous speculations about the cause of these variations. Most of these speculations start with a specific model and try to verify predictions particular to that model. The best known and most successful such approach starts from variations of the solar constant due to changes in the Earth's orbit^{2,3}.

Nicolis and Nicolis¹ have used a very different approach. Instead of using a specific model as a starting point, they asked whether the records of climatic change show the typical signs expected for any deterministic system. If these signs are not found, then one should conclude either that the system is randomly perturbed by some stochastic noise for which no deterministic description is possible, or that it is a deterministic system with many degrees of freedom. In the latter case, a deterministic description would be possible in principle, but rather unfeasible if the detailed mechanism were not known. If, in contrast, these signs were found, one should expect the system to be describable by a very simple model, with only few (≤ 4 in the present case, according to ref. 1) degrees of freedom. Note that the theory does not say anything about the details of such a model.

The method used in ref. 1 is based on recent developments in the theory of dynamical systems⁴. If a system develops deterministically, its trajectory typically stays on a low-dimensional submanifold of the total available phase space. Examples include fixed points and limit cycles. In addition to these, more and more 'strange attractors' have been found, for which the trajectory stays on a 'fractal'⁵ set, characterized by a non-integer dimension. Such systems are then deterministic only in the formal sense of being described by well-behaved differential equations, but they are neither periodic nor quasi-periodic (even if they are not disturbed irregularly), nor are they predictable in the long term⁶⁻⁸.

In searching for such behaviour, then, one can try to estimate the dimension of the phase-space region visited by the system. If we put points at random onto a d -dimensional object, we expect the number of points in any hypersphere of radius R to behave as R^d —provided that the centre of the hypersphere is also on the object. Thus, all one has to do is to count the average number of pairs in a time series $x_1, x_2, x_3, \dots$ (each x_i represents uniquely the state of the system at time $t = i\tau$, with some fixed delay τ) for which the two points are closer than a distance R .

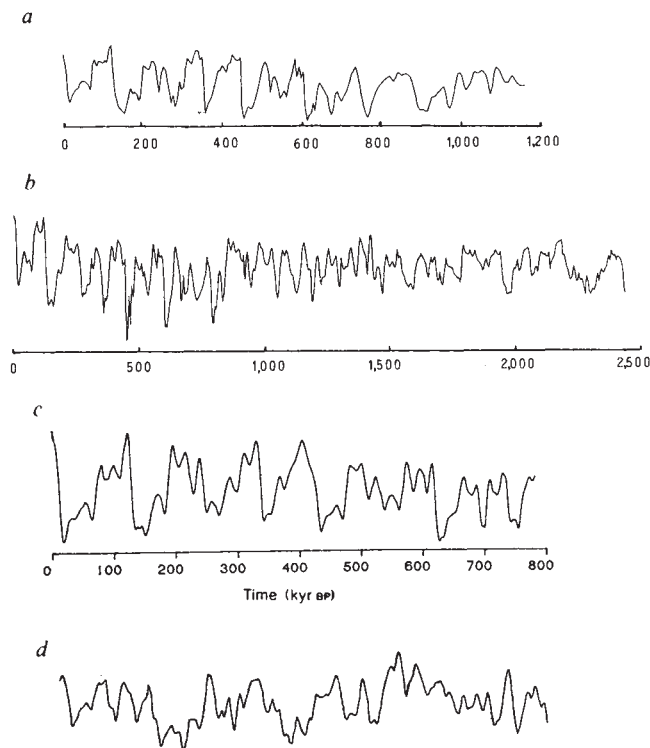


Fig. 1 $a-c$ ^{18}O variations (arbitrary scale) of different deep-sea cores. The data are smoothed, interpolated and plotted on a reconstructed timescale. *a*, Core V28-238, interpolation from ref. 16; *b*, core V28-239, interpolated from ref. 16; *c*, average over several cores, from ref. 15. *d*, Random time series, constructed and smoothed so as to simulate the statistical aspects of *c*.

If we find

$$N(R) \equiv (\text{number of pairs with } |\mathbf{x}_i - \mathbf{x}_j| < R) \sim R^d \quad (1)$$

we call d the (correlation) dimension of the attractor⁹.

In the present application, there is only one observable, ζ , measured at a single time. If the attractor has finite dimension, one can nevertheless construct the complete state vector $\mathbf{x}(t)$ by the following process^{10,11}: $\zeta(t+\tau)$ is used as the first coordinate, $\zeta(t+2\tau)$ as the second, ..., $\zeta(t+n\tau)$ as the last. If n is large enough, $\zeta_i = \{\zeta_{i+1}, \zeta_{i+2}, \dots, \zeta_{i+n}\}$ will be a faithful representation of $\mathbf{x}_i$. Different choices of n and of the delay τ will lead to differently shaped attractors, but all of them will have the same dimension, when using ζ_i instead of $\mathbf{x}_i$ in equation (1). A further simplification is to take the maximum norm in equation (1): we then obtain

$$N(R, n) \equiv (\text{number of pairs with } |\zeta_{i+k} - \zeta_{j+k}| < R; 0 < k \leq n) \sim R^d \quad (2)$$

Then we test equation (2) for increasing values of n , and investigate whether the effective value of d becomes independent of n and R .

Based on such an analysis, and using the ^{18}O data of the deep-sea core V28-238 (ref. 12) (for numerical raw values, see ref. 13), Nicolis and Nicolis¹ concluded that the worldwide climate of the past million years was restricted to an attractor of dimension $d \approx 3$. If correct, one should be able to write down closed equations for climatic changes with only four independent variables (although this gives no hint as to where these equations might come from and what they might look like).

To test this result, I have re-analysed the same data (with the proviso mentioned below), as well as data from other deep-sea cores (V28-239; ref. 14, and see below). I cannot confirm their conclusion. Instead, I find that the data are unable to give any upper bound for the dimension, while they yield a lower bound well above $d = 3$.

The main problems in applying equation (2) in the present case are: (1) errors in the estimate of the $^{18}\text{O}/^{16}\text{O}$ ratio (mainly due to local disturbances). These can be partly eliminated by smoothing and by averaging over several different cores¹⁵; (2)

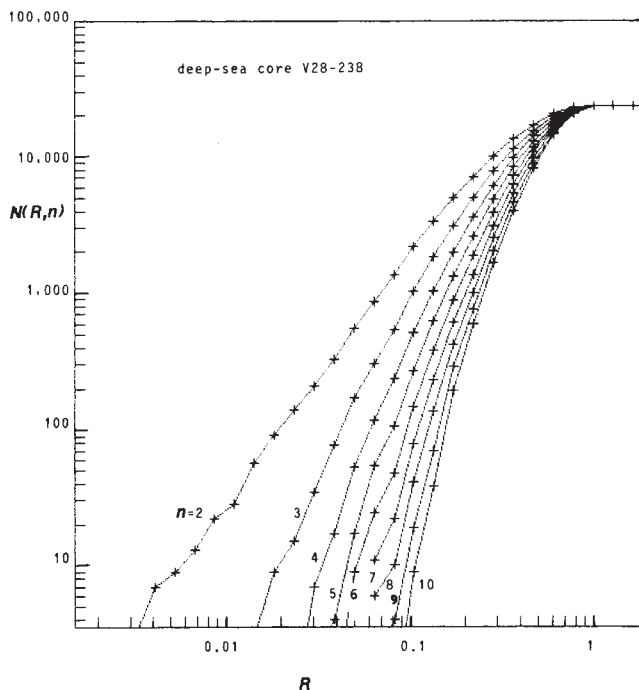


Fig. 2 Number $N(R, n)$ (see equation (2)) for deep-sea core V28-238, plotted against R on a doubly logarithmic scale. If variations of climate follow a strange attractor, the slopes would approach a constant d for $n \rightarrow \infty$. Data are from ref. 16.

uncertainties about the timescale. In a first attempt, one might assume that deposition on the sea floor occurs at a constant rate, but this is unrealistic^{15,16}. The most model-independent way of estimating timescales is to identify prominent events (such as extinctions of species or magnetic reversals). A less independent method involves the comparison of observed spectra with those expected from variations of the Earth's orbit¹⁵. Three curves obtained in this way are shown in Fig. 1a-c.

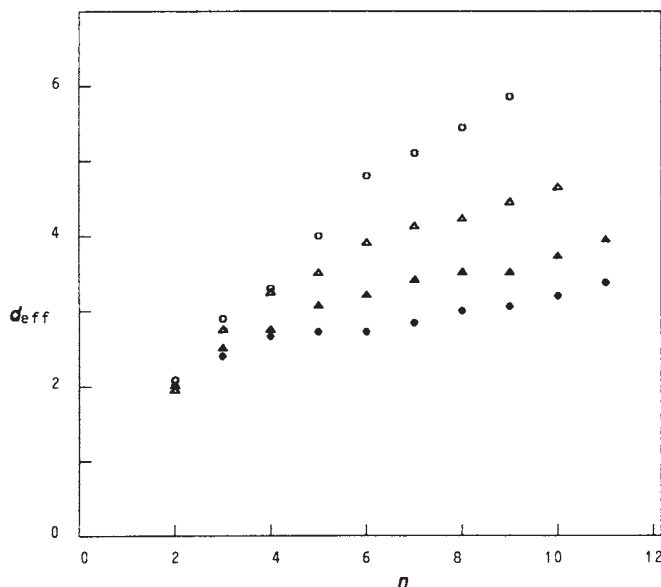


Fig. 3 Slopes from Fig. 2 (Δ) and from similar graphs using the data of Fig. 1b ($\circ$), 1c ($\bullet$) and 1d ($\blacktriangle$).

Figure 2 shows the number of pairs as in equation (2), using the data of core V28-238 as interpolated and processed in ref. 16 and shown in Fig. 1a. The proviso alluded to above is that Nicolis and Nicolis¹ and I have been provided by Pestiaux¹⁶ with different interpolations. While mine has 230 points, each separated by 5,000 yr, that used by Nicolis and Nicolis¹ has ~500 points separated by 2,000 yr. As the raw data consist of only 184 points, the 500 points used in ref. 1 are highly correlated, and should not be compared with random noise as is done there.

Equation (2) would require the data in Fig. 2 to fall on straight lines for sufficiently large fixed n , and for sufficiently small

values of R . Figure 2 shows no indication of that. Fitting straight lines in the region $20 < N(R, n) < 1,000$ gives the slopes shown in Fig. 3 (open triangles). According to equation (2), they should be equal to d for sufficiently large n . Again the data do not given any hint of a finite-dimensional attractor.

Figure 3 also shows (open circles) analogous results from core V28-239 (ref. 14), using the interpolation from ref. 16. These data, shown in Fig. 1b, extend over >2 Myr, and seem to be little disturbed¹⁶. Thus, they yield the clearest indication against a low-dimensional climatic attractor over the past 2 Myr.

The data shown in Figs 1c and 3 (dots) are from ref. 15. They represent averages over five different cores, the most important ones being V28-238, V22-174 and DSDP 502B. As shown in Fig. 1c, a rather severe filter was applied to these data. Also, they represent finer binning ($\tau = 2,000$ yr, as opposed to 5,000 yr in the first two data sets). It is thus not surprising that the effective values of d increase very slowly for these data. To test that such a slow increase of d is compatible with a random signal, we first constructed a random time series $\{y_i; i = 1, \dots, 400\}$ as $y_i = (5y_{i-1} + e_i)/6$, where e_i is uniformly drawn from the interval $[0, 1]$, and then applied a filter similar to that used in ref. 15. The resulting signal is shown in Fig. 1d. It is very similar to Fig. 1c, and yields similarly small dimension estimates (solid triangles in Fig. 3, obtained by averaging over five such runs). This is a warning that spuriously small dimension estimates can be obtained from using too few, too finely sampled and too highly smoothed data (note that the largest set of raw data in Fig. 1c, that of core V28-238, consists of only 184 points). Taking a small delay τ is not a disadvantage if it is feasible, and provided that the maximum norm is taken in equation (1). (Also, the sum in equation (1) should only include pairs (i, j) for which $|i - j|\tau$ is bigger than the correlation time.) Neglecting the latter has led to wrong conclusions (as in ref. 17). But one should not compare the results to unsmoothed random data, one should go to sufficiently large values of n , and one should beware of spurious plateaus as in Fig. 3.

I have also analysed tree-ring data from the White Mountains, California¹⁸, covering the past 7,100 yr. They are very different from the deep-sea core data: not only do they record a very local climate, but they cover a very different timescale. Thus, both data could *a priori* show completely different results. But again I have found a very high-dimensional attractor. Due to the vastly larger number of data points (7,100 compared with <500), this time the dimension can be estimated more seriously, with the result $d > 10$. (See Fig. 4).

Finally, the above analyses could also have provided estimated of Kolmogorov entropies^{7,8}, had d been finite. A small Kolmogorov entropy would be the most direct indication that the climate is predictable over the corresponding timescale. Unfortunately, because we have only bounds to the dimension, we can only obtain very weak lower bounds to the entropy. They are so weak that I do not want to quote them.

I thank Dr Berger for providing Pestiaux's interpolation of deep-sea cores V28-238 and V28-239, and Dr J. Robinson for sending me tree-ring data.

Received 17 March; accepted 3 July 1986.

1. Nicolis, C. & Nicolis, G. *Nature* **311**, 529-532 (1984); *Phys. Bull.* **41**, 5-9 (1985).
2. Milankovitch, M. *Théorie Mathématique de Phénomènes Thermiques Produits par la Radiation Solaire* (Gauthiers-Villars, Paris, 1920).
3. Berger, A. L. et al. (eds) *Milankovitch and Climate* (Reidel, Dordrecht, 1984).
4. Holden, A. V. (ed.) *Chaos—an Introduction* (Manchester University Press, 1985).
5. Mandelbrot, B. B. *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
6. Ruelle, D. *Math. Intell.* **2**, 126-137 (1980).
7. Shaw, R. Z. *Naturforsch.* **36**, 80-112 (1981).
8. Grassberger, P. in *Chaos—an Introduction* (ed. Holden, A. V.) 291-311 (Manchester University Press, 1985).
9. Grassberger, P. & Procaccia, I. *Physica* **9D**, 189-208 (1983); *Phys. Rev. Lett.* **50**, 346-349 (1983).
10. Packard, N. H., Crutchfield, J. P., Farmer, J. D. & Shaw, R. S. *Phys. Rev. Lett.* **45**, 712-716 (1980).
11. Takens, F. *Lect. Not. Math.* **898**, 366-381 (1981).
12. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55 (1973).

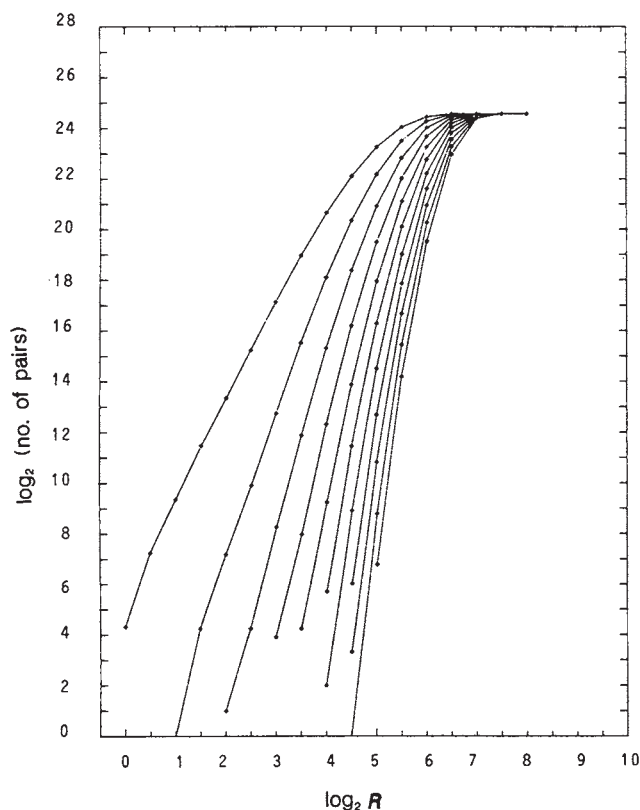


Fig. 4 As Fig. 2, but for tree-ring indices. Data from ref. 18.

13. Berger, A. L. & Pestiaux, P. *Tech. Rep.* No. 28 (Institute of Astronomy and Geophysics, Catholic University of Louvain, 1982).
14. Shackleton, N. J. & Opdyke, N. D. *Geol. Soc. Am. Mem.* 145, 449–464 (1976).
15. Imbrie, J. *et al.* in *Milankovitch and Climate* (eds Berger, A. L. *et al.*) Part 1, 269–305.
16. Pestiaux, P. *Analyse Spectrale de 14 Carottes Oceaniques du Quaternaire, Scientific Rep.* 1984/9 (Institut d'Astronomie et de Geophysique G. Lemaître, Université Catholique de Louvain-la-Neuve, 1984).
17. Fraedrich, K. J. *atmos. Sci.* (in the press).
18. Ferguson, C. W. *Tree-Ring Bull.* 29, (Laboratory of Tree-Ring Research, University of Arizona, Tucson, 1969).

Surface topography of (100)-type electro-faceted platinum from scanning tunnelling microscopy and electrochemistry

J. Gómez*, L. Vázquez*, A. M. Baró*, N. García*,
C. L. Perdriell†, W. E. Triaca† & A. J. Arvia†

* Departamento de Física Fundamental, Universidad Autónoma de Madrid, Canto Blanco, 28049-Madrid, Spain

† Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas, Universidad Nacional de La Plata, Sucursal 4, Casilla de Correo 16, 1900-La Plata, Argentina

The surface of a number of metals such as platinum, gold, rhodium and palladium, which are of particular interest in catalytic and electrocatalytic processes, either as polycrystalline or single-crystal specimens, can be preferentially orientated by a procedure known as electrochemical faceting^{1–3}. Here we report the first real-space observation in the nanometric range of electrochemically faceted (100)-type Pt electrodes, using scanning tunnelling microscopy (STM)⁴. Our findings confirm the existence of a faceted surface consisting of flat parts surrounded by ridge-type structures. At the initial stages cluster formation is observed, which subsequently develops in a preferential orientation. This constitutes direct evidence that faceting occurs through a selective electro-dissolution/electrodeposition process, which probably involves a fast diffusion of metallic species either in the electrolyte or on the surface.

The electrochemical faceting procedure consists of the application of a relatively fast periodic potential perturbation (Fig. 1a) to a metal electrode immersed in an acid solution in conditions suitable for developing a particular type of preferred orientation^{1,2} as inferred from SEM (scanning electron micrograph) patterns⁵ (Fig. 1b). Although there is indirect evidence of electrochemical faceting (Fig. 1c), it had not been observed at the nanometric range of surface micro-roughness. To follow the nature of the transformations associated with the different stages of faceting and its growth, it is important to have a real-space observation of the surface. At present, STM⁴ is the most suitable, non-destructive technique for studying the surface roughness of these Pt electrodes. This microscope has been successfully employed in the analysis of metal⁶ and semiconductor⁷ surface structures and the surface roughness of important technological materials^{8,9}, and in the determination of surface topography of biological specimens¹⁰. It provides a simultaneously high horizontal and vertical resolution offering three-dimensional real-space images of surfaces on an atomic scale, by working either under high vacuum or atmospheric pressure at room temperature.

The (100)-type preferentially oriented polycrystalline surfaces were made from polycrystalline Pt plates (~1 cm² apparent area) through the electrochemical faceting procedure, by applying a square-wave potential perturbation at a frequency $f = 2.5$ kHz between 1.4 and 0.05 V in 1 M H₂SO₄ at 25 °C. These conditions were maintained for up to 8 h. The degree of development of the (100)-type preferred orientation was followed by conventional voltammetry at 0.1 V s⁻¹ in the H-atom electroadsorp-

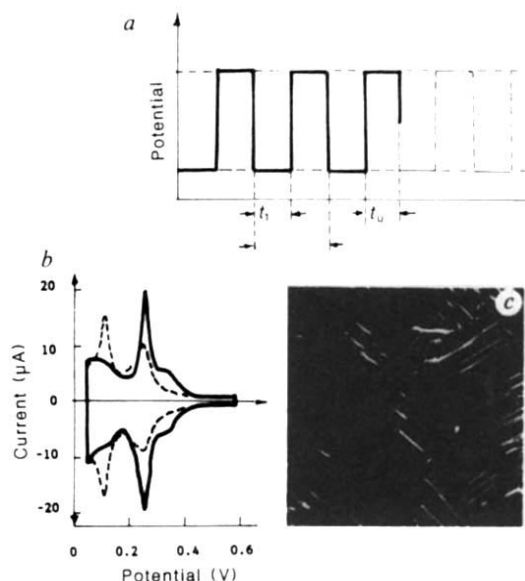


Fig. 1 *a*, Repetitive square-wave potential programme used to develop the (100)-type preferentially orientated Pt surface from polycrystalline Pt. *b*, Voltammograms with Pt electrodes at 0.1 V s⁻¹ in 1 M H₂SO₄ at 25 °C, covering the H-atom electroadsorption/electrodesorption potential range. Dashed line, untreated mirror-polished polycrystalline Pt electrode (1 cm² geometric area). Solid trace indicates the voltammogram (third cycle) of the (100)-type preferentially orientated polycrystalline Pt resulting after 8 h of electrochemical faceting in the conditions indicated in *a*. A clear increase in the charge of the strongly bound H-atom and a corresponding decrease in the charge of the weakly bound H-atom electroadsorption/electrodesorption peaks are observed. The voltammetric charge resulting from both electrodes is practically the same. This indicates that, in this case, the electrochemical faceting implied no change in the electrode roughness. *c*, SEM pattern of a (100)-type preferentially orientated polycrystalline Pt surface resulting from 8 h of electrochemical faceting as described in *a*. Parallel channel-like structures are observed. Scale bar, 10 μm.

tion/electrodesorption potential range^{1,2}. The treated Pt electrodes were removed from the electrochemical cell, rinsed in distilled water and mounted in the STM sample holder. STM measurements were made at room temperature and 10⁻⁶ torr. Each specimen was systematically analysed at different areas, separated by macroscopic distances.

An STM pattern (13,000 × 11,700 Å) of polycrystalline Pt which has been mirror-polished without electrochemical faceting (blank) (Fig. 2a) shows an inhomogeneous surface topography with the absence of any preferential orientation. The r.m.s. average surface roughness is calculated to be 110 Å, as expected for very flat, mechanically polished metal specimens. An STM pattern (14,000 × 6,000 Å) (Fig. 2b) after 2 h of electrochemical faceting in the experimental conditions described above reveals a surface smoother than that of the blank, with a few steps of variable length (*x*-axis) and of ~100 Å height (*z*-axis). In addition, the treated surface shows cluster formations, which in most cases are well localized in space (Fig. 2b). At this stage, the average height and width of clusters is estimated at ~100 Å and ~500 Å. In certain regions (Fig. 2c) the clusters show a definite orientation, forming ridges. For a longer electrochemical treatment, the development of ridges furnishes large flat plateaux and grooves (Fig. 3a, after an 8-h treatment). Figure 3a (11,700 × 12,000 Å) shows the prevalence of parallel ridges ranging from ~50 Å to ~100 Å in height.

Other surface regions of the same specimen (Fig. 3b) show a dramatic development of smooth, square plaquettes of relatively large area (3,000 Å²), terminating in pronounced grooves. Ridge-like structures in two directions at nearly 90° are observed,

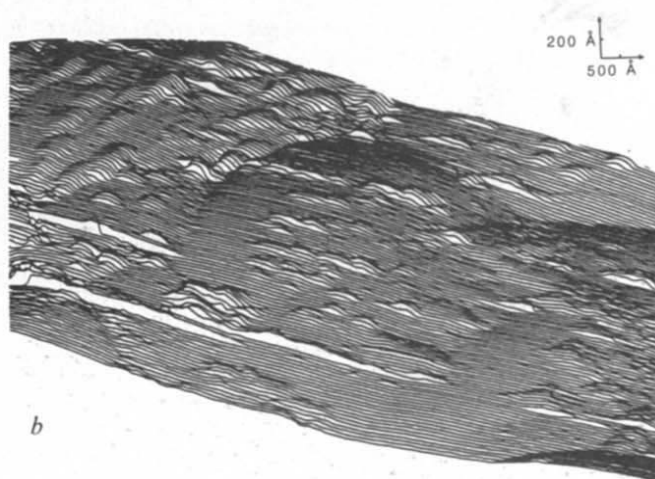
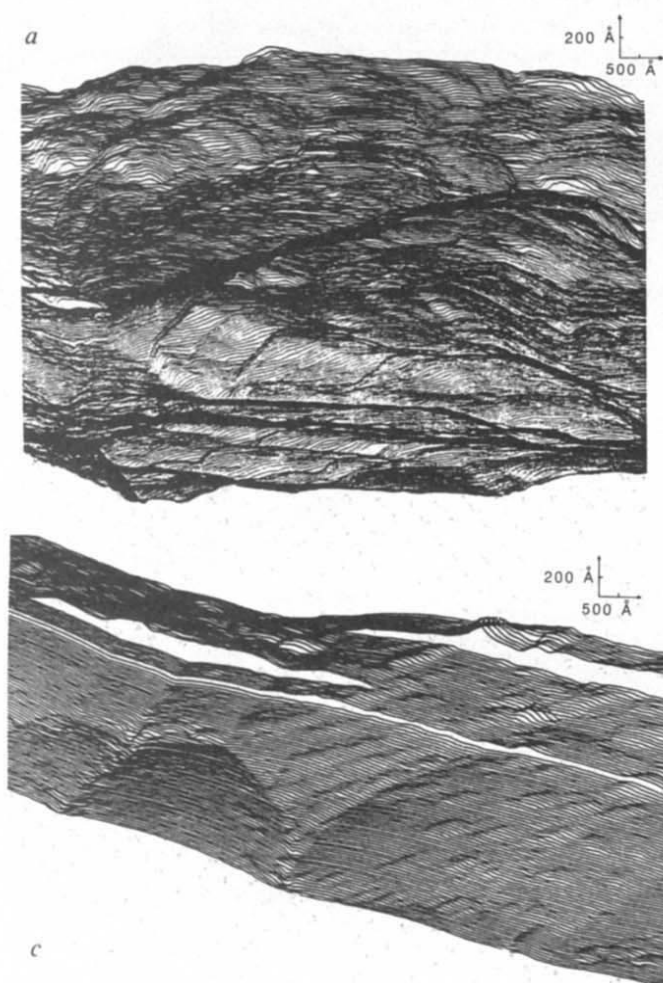


Fig. 2 *a*, STM pattern of a mirror-polished untreated polycrystalline Pt (blank) obtained under high vacuum (10^{-6} torr) and room temperature. An inhomogeneous corrugation without any preferential orientation is observed. The average surface roughness is ~ 200 Å. The surface sampling corresponds to 255 lines covering $17,000$ Å along the y -axis. $\Delta y = 11,700$ Å. *b*, STM pattern of a (100)-type faceted polycrystalline Pt after 2 h of electrochemical faceting in the conditions described in Fig. 1*a*. STM micrography was obtained in the same conditions indicated for the blank. In comparison with the STM pattern in *a* for the untreated surface, one observes a relatively smooth surface, with a few steps randomly distributed parallel to the scanning direction, with an average height of ~ 100 Å. Flat clusters and ridge-type configurations are distributed in different surface regions with a typical height of 100 Å and width of 500 Å. Surface sampling corresponds to $14,000 \times 6,500$ Å. $\Delta y = 6,050$ Å. *c*, STM pattern of another region from the specimen in *a*, showing large steps parallel to the scanning direction and the development of clusters along a preferential direction nearly perpendicular to the step direction. $\Delta y = 5,900$ Å.

related to the symmetry of the (100) surface. This ridge-like topography is similar to that reported for Au (100) single crystals in air at atmospheric pressure⁸.

Our STM results support the interpretation of the propagation of electrochemical faceting based on the selective cyclic electro-dissolution and electrodeposition of Pt in the acid electrolyte produced during the fast, periodic potential perturbation treat-

ment³. In these conditions, the average thickness of the apparent diffusion boundary layer (δ_N) decreases with the square root of the period of the perturbation¹¹, so that for frequencies of the order of 1 kHz, and a diffusion coefficient of soluble ionic Pt of $\sim 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, δ_N is of the order of 10^{-5} cm. This distance corresponds to the estimated r.m.s. displacement of dissolved ionic Pt at the same frequency. As this average diffusion layer

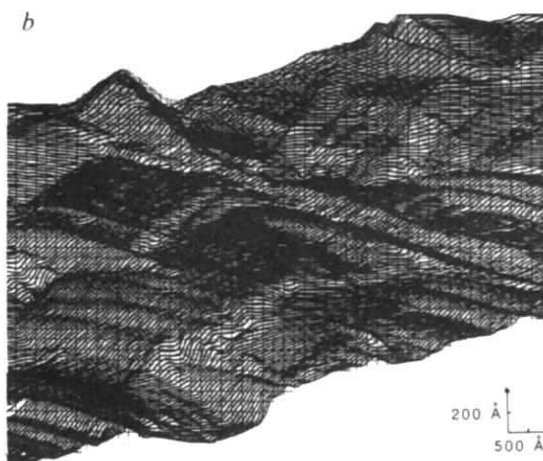
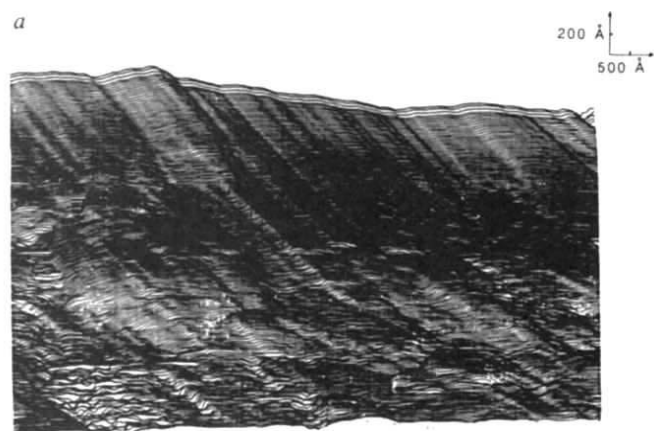


Fig. 3 *a*, STM pattern of a (100)-type faceted polycrystalline Pt after 8 h of electrochemical faceting in the conditions described in Fig. 1*a*. A parallel-ridge-like structure prevails over most of the surface that may be a remnant of earlier stages. In the lower part, flat clusters are still observed. In the upper part, steps of the order of 10 Å can be distinguished. The average ridge extension along the y -axis is $\sim 5,000$ Å and the ridge separation, although variable, lies between $1,200$ and 200 Å. The ridge height ranges from ~ 50 to 100 Å. $\Delta y = 12,000$ Å. *b*, STM pattern of (100)-type faceted polycrystalline Pt, from the same specimen shown in *a*, taken in a region separated by several micrometres from that shown in *a*. Ridge-like structures in two directions at nearly 90° are observed. Large plateaux surrounded by ridges are also seen, with a roughness of < 20 Å. The maximum local height of surface corrugation is ~ 300 Å. An outstanding formation of plaquettes is seen. $\Delta y = 6,500$ Å.

becomes smaller than the surface irregularities, all points at the electrode surface are equally accessible for the electrodeposition ions. This implies an electrodeposition without roughness development, as deduced from the constant value of the total voltammetric charge resulting from integration of the current potential profile before and after the electrochemical faceting (Fig. 1b).

We believe that the sequence of events is: (1) during each anodic half-cycle, Pt is selectively electrodisolved (from sites where Pt atoms are weakly bound) by the magnitude of the upper limit of the potential perturbation, therefore selecting the surface that will be preferentially reorientated, namely the (110), (100) and (111) surfaces³; (2) during each reversed half-cycle, Pt atoms are selectively electrodeposited at surface sites exhibiting a stronger metal-metal bond energy. These results indicate that the local and selective deposition/dissolution sequence proceeds through the formation of clusters at the earlier stages of the electrochemical treatment (Fig. 2b), yielding later ridge-like structures (Fig. 3a) of practically the same characteristics as those grown by anodic deposition⁸. The metal-atom diffusion rate at the metal-electrolyte interface should be much larger than at the metal-metal gas interface. This is because Pt deposited from the gas phase on a flat Pt surface grows, at room temperature, in a pyramid structure as we have measured by STM⁸, instead of in flat clusters and ridges as observed in electrochemical faceting.

We conclude that the propagation of electrochemical faceting comprises initially the appearance of nuclei which develop into

a preferentially orientated ridge-like surface topography. This means that ridges represent a stable configuration of this surface. Note that the development of (100)-type preferentially orientated Pt surface occurs without any STM-measurable change in roughness and that the (100)-type preferentially orientated Pt resulting from the electrochemical faceting is very stable, exhibiting the same topography in air when observed by STM and by voltammetry. This high stability is probably a consequence of the severe electrochemical treatment.

We thank J. Soler, H. Rohrer, M. Aguilar, J. Presa, J. Pedrosa, J. Gonzalo-Velasco and Marisa Marcos, for discussions and the Centro Científico UAM-IBM, CAI, CYT through contracts 1426/82 and 0386/84 and Fundación Ramon Areces for financial support.

A.J.A. is Visiting Professor, Departamento de Electroquímica and Física Fundamental, Universidad Autónoma de Madrid.

Received 11 February; accepted 13 June 1986.

- Canullo, J. C., Triaca, W. E. & Arvia, A. J. *J. electroanal. Chem.* **175**, 337-348 (1984).
- Cervino, R. M., Triaca, W. E. & Arvia, A. J. *J. electrochem. Soc.* **132**, 266-280 (1985).
- Arvia, A. J., Canullo, J. C., Custidiano, E., Perdiel, C. L. & Tiraca, W. E. *Electrochim. Acta* (in the press).
- Binnig, G., Rohrer, H., Gerber, Ch. & Weibel, E. *Appl. Phys. Lett.* **40**, 178-181 (1982).
- Cervino, R. M., Arvia, A. J. & Vielstich, W. E. *Surf. Sci.* **154**, 623-630 (1985).
- Binnig, G. & Rohrer, H. *Helv. phys. Acta* **55**, 726-630 (1982); *Surf. Sci.* **126**, 236-265 (1983).
- Binnig, G., Rohrer, H., Gerber, Ch. & Weibel, E. *Phys. Rev. Lett.* **50**, 120-123 (1983).
- Miranda, R. *et al. Appl. Phys. Lett.* **47**, 367-369 (1985).
- Garcia, N. *et al. Metrologia* **21**, 135-138 (1985).
- Baró, A. M. *et al. Nature* **315**, 253-254 (1985).
- Conway, B. E., Bockris, J. O'M., Yeager, E., Khan, S. U. & White, R. E. (eds) *Comprehensive Treatise of Electrochemistry* Vol. 7, Chs. 5, 6, (Plenum, New York, 1983).

Room-temperature ionic liquids as solvents for electronic absorption spectroscopy of halide complexes

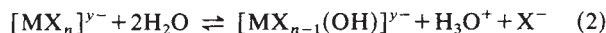
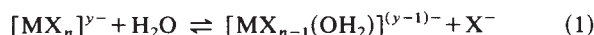
Denise Appleby*, Charles L. Hussey†, Kenneth R. Seddon* & Janet E. Turp*

* School of Chemistry and Molecular Sciences, University of Sussex, Falmer, Brighton BN1 9QJ, UK

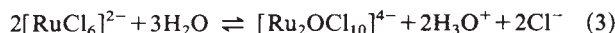
† Department of Chemistry, University of Mississippi, University, Mississippi 38677, USA

Solution studies of halide complexes of the transition metals, of general type $[MX_n]^{y-}$ ($X = Cl$ or Br), have been severely hampered by the existence of both solvation and solvolysis phenomena¹. We demonstrate here that the use of room-temperature ionic liquids (often misleadingly referred to as molten salts) as solvents circumvents both of these problems, and permits the first reliable solution spectra to be recorded for these species; in some cases the spectra exhibit quite remarkable resolution.

In water, the following hydration and hydrolysis equilibria exist:



In a number of well-documented cases, the hydrolysis reaction results in the formation of oxo-bridged dimers²; for example,



The use of solvents such as ethanenitrile (MeCN, where 'Me' stands for the methyl group, CH_3) and dimethyl sulphoxide (Me_2SO), in which solvolytic reactions are unlikely, usually results in an enhanced solvation problem, whereas solvents of low dielectric constant (such as dichloromethane) often produce spectral distortion due to the formation of discrete, intimate ion pairs. Thus, many of the electronic absorption spectra reported

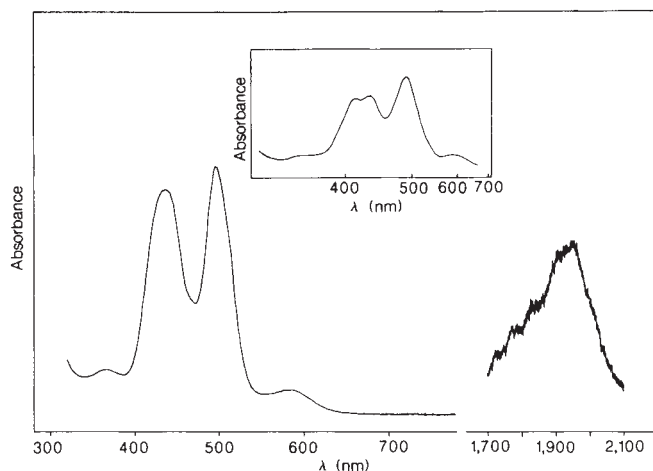


Fig. 1 Visible and near-infrared spectra of $[MeEtim][IrCl_6]$ in a basic $[MeEtim]Cl/AlCl_3$ ionic liquid; the absorbance scale in the near-infrared region is 10^3 times more expanded than that in the high-energy region. The inset is the classic absorption spectrum¹⁰ of the hexachloroiridate(IV) anion in concentrated hydrochloric acid.

in the literature as the spectra of $[MX_n]^{y-}$ are often either poorly defined or significantly distorted, due to the presence of solvation and/or solvolysis products, a problem openly recognized and discussed by Jørgensen as long ago as 1963³.

High-temperature molten salts based on halides (such as $LiCl/KCl$ or $NaCl/AlCl_3$) have been used to study the spectroscopic properties of $[MX_n]^{y-}$ for many years, but the classic work of Gruen and McBeth⁴ identifies the following problems with the spectra thus obtained: (1) distortion due to thermal broadening; (2) thermally promoted dissociation reactions, such as $[VCl_6]^{3-} \rightleftharpoons [VCl_4]^- + 2Cl^-$; and (3) thermally promoted disproportionation reactions, such as $2[MoCl_6]^{2-} \rightleftharpoons [MoCl_6]^- + [MoCl_6]^{3-}$. Although room-temperature, halide-based ionic

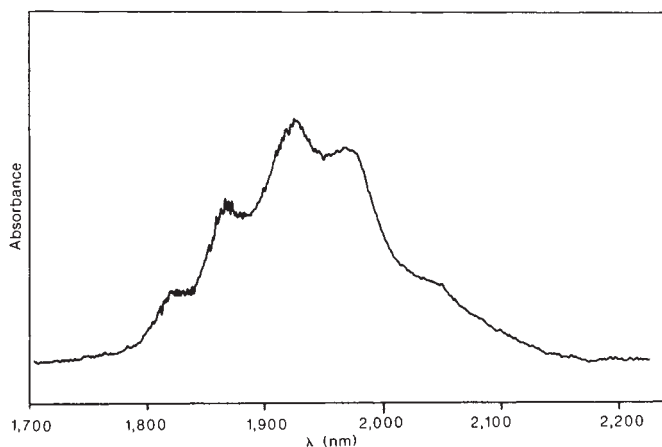


Fig. 2 Near-infrared spectrum of $[\text{MeEtim}]_2[\text{OsCl}_6]$ in a basic $[\text{MeEtim}]\text{Cl}/\text{AlCl}_3$ ionic liquid.

liquids have been known since 1952⁵, their potential use as solvents for spectroscopy has never been explored beyond their use for the 'fingerprinting' of the tetrachlorocobaltate(II) anion in an *N*-butylpyridinium chloride:aluminium chloride ($[\text{NBupy}]\text{Cl}/\text{AlCl}_3$) ionic liquid⁶. We present here a selection of spectra which have been obtained using ionic liquids based on 1-methyl-3-ethylimidazolium chloride:aluminium chloride ($[\text{MeEtim}]\text{Cl}/\text{AlCl}_3$)⁷ and 1-methyl-3-ethylimidazolium bromide:aluminium bromide ($[\text{MeEtim}]\text{Br}/\text{AlBr}_3$)⁸ mixtures, in order to illustrate their wide applicability for the study of halide complexes. All of the spectra discussed here were obtained in 'basic' compositions of the ionic liquid (those with a ratio of $[\text{MeEtim}]\text{X}:\text{AlX}_3 > 1:1$)⁹, which ensures the presence of a large excess of X^- ; identical results were obtained in basic $[\text{NBupy}]\text{Cl}/\text{AlCl}_3$ ionic liquids. The chemical nature of the solvent precludes the occurrence of both solvation reactions (equation (1)) and solvolysis reactions (equation (2)). Adventitious molecular water (ubiquitous in conventional molecular solvents) is destroyed by reaction with the chloroaluminate anions.

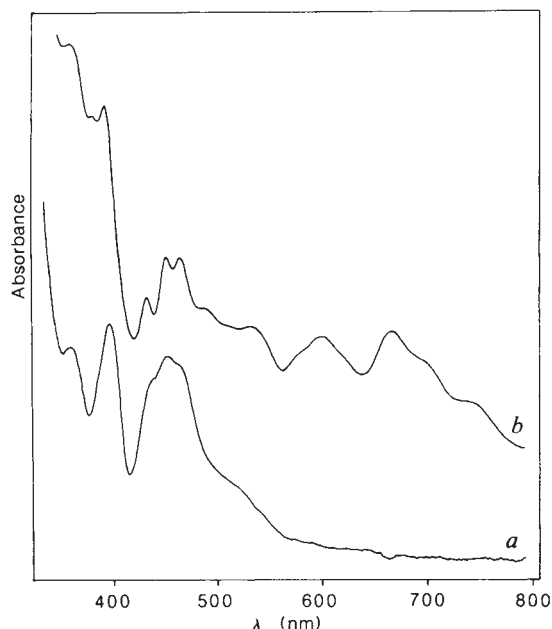


Fig. 3 Electronic spectra of $[\text{pyH}]_2[\text{RuBr}_6]$ in 6 M hydrobromic acid (a) and a basic $[\text{MeEtim}]\text{Br}/\text{AlBr}_3$ ionic liquid (b).

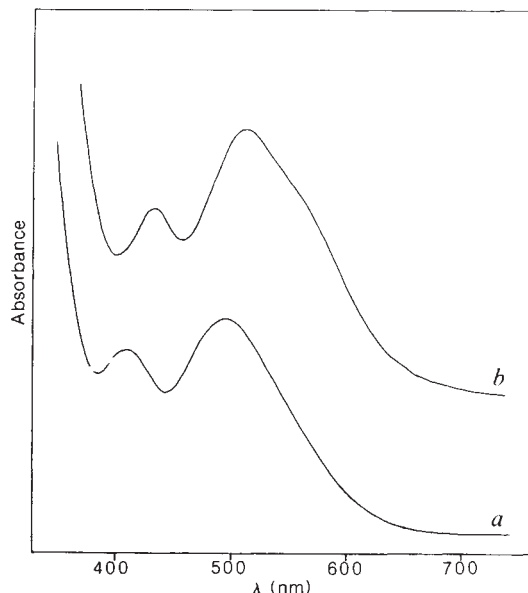


Fig. 4 Electronic spectra of $[\text{MeEtim}]_3[\text{Ru}_2\text{Br}_9]$ in 6 M hydrobromic acid (a) and a basic $[\text{MeEtim}]\text{Br}/\text{AlBr}_3$ ionic liquid (b).

Figure 1 shows the visible and near-infrared spectra of the hexachloroiridate(IV) anion, $[\text{IrCl}_6]^{2-}$. The spectral shifts and profile changes for the visible region, compared with the classical spectrum presented by Jørgensen (Fig. 1 inset)¹⁰, exemplify the significant distortions present due to reactions of types (1) and (2), even in concentrated hydrochloric acid. Moreover, a band is observed in the near-infrared region corresponding to a transition between the two spin-orbit components of the $^2T_{2g}$ electronic ground state ($E_g'' \rightarrow U_g'$). This is the first recorded observation of this band for a solution species, although the presence of such a transition has been inferred from electronic resonance Raman spectroscopy¹¹ and observed by diffuse reflectance measurements performed on the solid¹². Moreover, the splitting of this band into two components demonstrates unequivocally the presence of the disputed¹¹ Jahn-Teller distortion of the U_g' state: its 4-fold degeneracy is lifted to produce two Kramers doublets, E_g'' and E_g' . Similarly, the observation of intra-configurational ($^3T_{1g}$) transitions^{11,13} for $[\text{OsCl}_6]^{2-}$ in the basic $[\text{MeEtim}]\text{Cl}/\text{AlCl}_3$ ionic liquid (see Fig. 2) is also reported here for the first time. The importance of observing spectroscopic properties in an isotropic homogeneous medium cannot be overstated: it eliminates the ambiguity of transitions which may be induced by reduced site symmetry in the solid state.

The recent development of the first room-temperature, bromide-based ionic liquid, 1-methyl-3-ethylimidazolium bromide:aluminium bromide ($[\text{MeEtim}]\text{Br}/\text{AlBr}_3$)⁸, has enabled the first true solution electronic spectra to be obtained for bromide complexes. This is demonstrated by comparison of the solution electronic spectra of $[\text{pyH}]_2[\text{RuBr}_6]$ (Fig. 3; py = pyridine) and $[\text{MeEtim}]_3[\text{Ru}_2\text{Br}_9]$ ¹⁴ (Fig. 4) in both 6 M hydrobromic acid and a basic bromide ionic liquid. For the former complex, spectral resolution approaching that obtained at $<14\text{ K}$ in the solid state^{15,16} is observed in the ionic liquid, and for the latter salt, large shifts are observed in both the positions and relative intensities of the bands. We have also studied, in both the chloride and bromide systems, the simple halide complexes of titanium, vanadium, chromium, manganese, iron, cobalt, nickel, copper, molybdenum, tungsten, rhenium, ruthenium, osmium, rhodium and iridium. In all cases, we observe enhanced resolution and significant spectral shifts compared with the standard published solution spectra; our results most closely resemble those recorded in solvents of low dielectric constant (such as dichloromethane).

We thank the SERC for the award of a postdoctoral fellowship

(J.E.T.) and financial support. Work performed at the University of Mississippi was supported by NSF grant CHE-8412730 and funds for collaborative research were provided by NATO grant 020/84.

Received 5 March; accepted 16 June 1986.

1. Lever, A. B. P. *Inorganic Electronic Spectroscopy* 2nd edn (Elsevier, Amsterdam, 1984).
2. Seddon, E. A. & Seddon, K. R. *The Chemistry of Ruthenium* (Elsevier, Amsterdam, 1984).
3. Jørgensen, C. K. *Inorganic Complexes* (Academic, London, 1963).
4. Gruen, D. M. & McBeth, R. L. *Pure appl. Chem.* **6**, 23-47 (1963).
5. Hurley, F. H. & Wier, T. P. *J. electrochem. Soc.* **98**, 203-206 (1951).
6. Laher, T. M. & Hussey, C. L. *Inorg. Chem.* **20**, 4201-4206 (1981).
7. Wilkes, J. S., Levisky, J. A., Wilson, R. A. & Hussey, C. L. *Inorg. Chem.* **21**, 1263-1264 (1982).
8. Sanders, J. R., Ward, E. H. & Hussey, C. L. *J. electrochem. Soc.* **133**, 325-330 (1986).
9. Hussey, C. L. *Adv. molten Salt Chem.* **5**, 185-229 (1983).
10. Jørgensen, C. K. *Molec. Phys.* **2**, 309-332 (1959).
11. Clark, R. J. H. & Turtle, P. C. *Chem. Phys. Lett.* **51**, 265-268 (1977); *JCS Faraday Trans. II* **74**, 2063-2076 (1978).
12. Allen, G. C., Al Mobarak, R., El-Sharkawy, G. A. M. & Warren, K. D. *Inorg. Chem.* **11**, 787-796 (1972).
13. Ikeda, K.-I. & Maeda, S. *Inorg. Chem.* **17**, 2698-2701 (1978).
14. Appleby, D. et al. *JCS chem. Commun.*, 483-485 (1986).
15. Clark, R. J. H. & Dines, T. J. *Molec. Phys.* **52**, 859-870 (1984).
16. Collingwood, J. C., Schatz, P. N. & McCarthy, P. J. *Molec. Phys.* **30**, 469-491 (1975).

Basal reversals in layered intrusions are evidence for emplacement of compositionally stratified magma

J. R. Wilson & O. Engell-Sørensen

Geologisk Institut, C. F. Møllers Alle, 8000 Aarhus C, Denmark

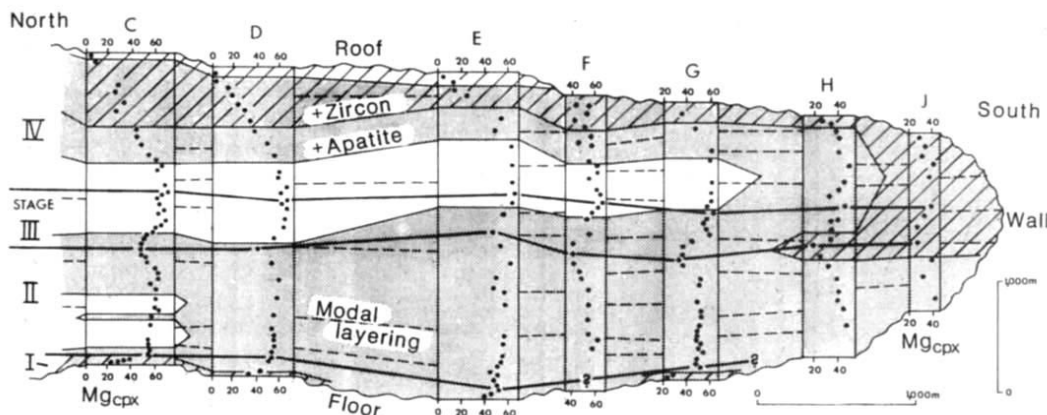
Compositional reversals, with rocks becoming more primitive upwards, is a feature at the base of many layered intrusions, such as the Stillwater Complex¹, the Muskox intrusion² and the Great Dyke³. Mechanisms such as contamination by an evolved melt from partial anatexis of country rocks or reaction between cumulus minerals and intercumulus melt are usually inadequate to explain these features⁴. A major reversal at the floor of the Hyllingen Series, part of the Fongen-Hyllingen complex, Norway, is consistent with crystallization during gradual elevation of compositionally stratified magma up an inclined surface in response to the influx of dense, primitive magma. During enlargement of the chamber, increasingly primitive magma comes into contact with the sloping floor just below the roof and an inverted sequence of crystalline products derived from the stratified magma develops. An intrusive mechanism of this type seems likely for layered intrusions which possess basal Mg-enrichment trends.

The Caledonian Fongen-Hyllingen complex, Norway (11.5° E, 63° N), is a major (160 km²) synorogenic layered mafic intrusion⁵. A wide range of fractionation products crystallized in elevated pH₂O conditions from a broadly tholeiitic parent magma. A thickness of ~10 km of layered rocks occur, and periodic reversals in cryptic variations indicate repeated addition of new magma. The northern (Fongen) part of the complex represents the lower part of the intrusion while the southern area (Hyllingen) represents the upper part with quartz-bearing syenitic rocks at the roof in contact with country rock amphibolites. Modal layering in most of the Hyllingen Series strikes roughly north-south, directly towards the southern margin, and dips eastward at ~40°, exposing a continuous sequence from floor to roof. All the Hyllingen Series rocks are relatively evolved. Diorites and ferrodiorites dominate. Wilson and Larsen⁶ subdivided the Hyllingen Series into four stages on the basis of vertical compositional evolution patterns (Fig. 1). Stage I (50-200 m) at the base, has highly evolved compositions. This is followed by stage II (600-1,400 m) which has fairly constant compositions, except at the top, where the rocks become progressively evolved. Stage III (325-500 m) is characterized by a gradual regression. Stage IV covers a wide fractionation range, reaching extremely evolved compositions at the roof. A vital feature is the lateral compositional variation. For example, approaching the southern margin along the stage III-IV boundary over a distance of ~5 km, Mg_{cpx} (=100 Mg/(Mg+Fe) in Ca-rich pyroxene) values fall systematically from 73 to 38.

Wilson and Larsen⁶ explained these features by crystallization of compositionally stratified magma along an inclined surface, where the plane of modal layering represents the crystallization front and cryptic layering reflects magma stratification (Fig. 2). A recent study of the Bushveld Intrusion⁷ has revealed vertical and lateral variations which can be explained by this model. The concept is extended to explain the basal reversal in the Hyllingen Series described below.

The marginal diorite at the base of the Hyllingen Series is in contact with basic metavolcanics (Fig. 3), while metapelites form the floor farther north^{5,8}. A tongue of diorite protrudes southwards into the country rocks at the base of profile B, indicating that intrusion was from a source lying to the north. The marginal diorite develops intermittent modal layering 40-220 m from the base and there is a discordance of about 7° between modal layering and the marginal diorite. Profile C has highly evolved compositions at the base and a systematic upward Mg-enrichment trend from Mg_{cpx} 22 to 57. After this regression the compositions remain fairly constant through >800 m of layered rocks. The compositions at the base of profile C are equivalent to those 2,400 m stratigraphically above (Fig. 1). This trend is, to some

Fig. 1 Correlation of stratigraphic profiles in the Hyllingen Series, simplified after ref. 6 and including some of the new data from Fig. 3. Profile numbers are the same as in ref. 6. Mg_{cpx} = 100 Mg/(Mg+Fe) in Ca-rich pyroxene. The stage boundaries are defined thus: I/II and III/IV—Mg_{cpx} values cease to increase systematically upwards; II/III—Mg_{cpx} values start to increase systematically upwards. The compositional regression of stage III is interpreted as reflecting crystallization during the gradual influx of new magma. See Fig. 2.



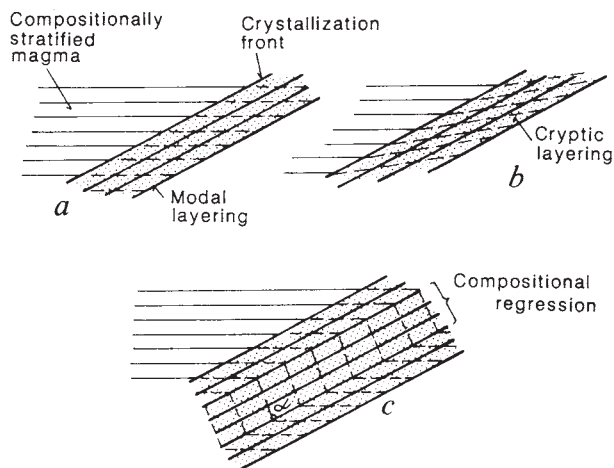


Fig. 2 *a*, Development of discordant relations between modal and cryptic layering by crystallization of stratified magma along an inclined front. *b*, Cryptic layering, reflecting the compositional stratification of the magma, may dip at an angle intermediate between that of the crystallization front and the horizontal. *c*, Compositional regression formed by crystallization along an inclined floor during uplift of stratified magma in response to influx of dense, primitive magma. The angle α will depend on the relative rates of crystallization and magma influx.

degree, observed in all the profiles. Profile E is the least convincing in this respect but the lowest sample is ~ 70 m above the floor. In profile C the entire regression is developed in the lowest 60 m. Above the basal regression there is a systematic variation towards more evolved compositions from north to south, from an average Mg_{cpx} value of 57.8 in profile C to 49.4 in profile G (Fig. 3).

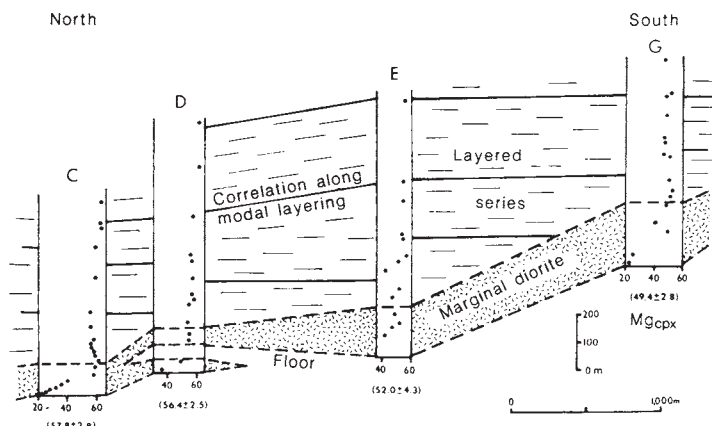
Incorporation of decreasing amounts of evolved material derived by partial melting of country rock metapelite cannot conceivably produce the observed Fe:Mg ratios, the systematic nature of the regression, or explain the abundant presence of apatite and zircon. Reaction between cumulus minerals and intercumulus melt is also inadequate to explain the observed features. The basal reversal is believed to result from crystallization along an inclined floor during the elevation of compositionally stratified magma (Fig. 4). The chamber is envisaged as having a wedge shape which inflates in response to the influx of dense primitive magma at the base. Evolved buoyant magma would first come into contact with the floor at the leading edge of the expanding wedge where some crystallization would occur. During continued magma chamber expansion, successively more primitive magma would come into contact with the floor where the crystallization sequence would be the reverse of that in the parental stratified magma. During formation of the reversal a large proportion of the heat loss would be through

the floor, but as the thickness of the basal rocks increases and the floor heats up, most of the heat would be lost through the roof. The thickness of the basal reversal depends largely on the thermal gradient between the floor and the magma and on the rate of elevation of the stratified magma. The compositional range of the reversal would depend on that of the parental stratified magma. Modal layering would start to develop at some distance from the floor when the cooling rate decreases sufficiently to allow oscillatory nucleation in slightly supercooled magma⁹. A stratified magma chamber (Fongen) was already present to the north before development of the Hyllingen Series, and this is envisaged as having expanded southwards by the process outlined above. The systematic nature of the regression in profile C (Fig. 3) indicates that the stratification of the magma was gradational with no compositional steps, and that influx occurred continuously over a long interval. Influx by a series of pulses may be implied by the irregular nature of the basal reversals in the other profiles. The fairly constant compositions above the basal reversal reflect crystallization during elevation of the stratified magma, minor reversals and normal fractionation sequences reflecting relative rates of crystallization and magma addition (Fig. 2c). The lateral compositional variation in the layered rocks is consistent with the crystallization of stratified magma along an inclined floor (Fig. 2). Crystallization along this inclined floor released buoyant, depleted magma which migrated up the slope to replenish the overlying, evolved magma layers¹⁰.

Several layered intrusions have compositional reversals at the base. The Stillwater Complex has a basal Mg-enrichment trend with Mg_{opx} ranging from 76 to 86 (the basal norite has Mg_{opx} 70)¹, overlain by 2,000 m of layered rocks with Mg_{opx} usually in the range 87–84. Compositions equivalent to those at the base do not occur until well within the banded zone. Similarly the Hartley Complex of the Great Dyke has a norite at the base with Mg_{opx} in the range of 81–84, followed by a small but systematic reversal from Mg_{opx} 87.5 to 89 below a thick sequence of ultramafic rocks with compositions more primitive than Mg_{opx} 89 (ref. 3). Compositions equivalent to those at the base do not occur until a stratigraphic height of ~1,700 m. The Muskox intrusion also has a marginal zone (120–200 m thick) which shows an upward Mg-enrichment from Fe_{70} to Fe_{85} , followed by a thick sequence of layered ultramafic rocks in which olivine is generally in the range Fe_{80-90} (refs 2, 11). Gabbroic rocks equivalent to those in the basal reversal are not encountered until ~1,800 m higher stratigraphically. Furthermore, the chilled margin which occurs sporadically at the floor of the Muskox marginal zone is considerably more evolved than the estimated primitive parental liquid¹¹.

Raedeke and McCallum¹ considered that reaction between cumulus minerals and varying amounts of intercumulus melt was the most likely mechanism to explain the Stillwater basal reversal, although they calculated that this could not account

Fig. 3 Correlation of stratigraphic profiles from the lower part of the Hyllingen Series, simplified after Wilson *et al.*¹⁵. The column widths depend on the range of Mg_{cpx} values present. Profiles are arranged so that the plane of modal layering is approximately horizontal. Profile numbers are the same as in Fig. 1. The numbers in brackets below each column refer to the mean and standard deviation of Mg_{cpx} values for the relevant profile above the basal regression. The average sample interval is 46 m, ranging from 33 m (profile C) to 62 m (D). Along the boundary between the layered series and the marginal diorite a thickness of ~600 m of layered rocks wedge out over a distance of ~4 km, a discordance of about 7°. The compositional regressions are also shown by other minerals. For example, the profile c regression has olivine (Fo₄₋₃₀), plagioclase (An₃₅₋₄₇), calcic amphibole (Mg_{hbl} 15-40), accompanied by the successive upward disappearance of zircon and apatite (Fig. 1).



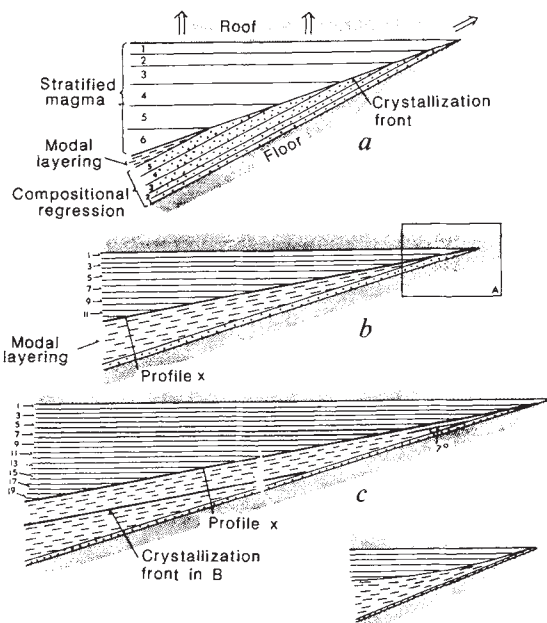


Fig. 4 Formation of a compositional reversal at the base of a layered intrusion by elevation of stratified magma along an inclined floor. *a*, Detailed relations in the wedge-shaped protrusion of an expanding magma chamber. The compositional sequence of the solid product on the floor will be the reverse of the stratified parental magma. *b*, Possible relationship between modal layering, stratified magma and the geometry of the expanding magma chamber. Magma layers 1–6 in *a* are equivalent to magma layers 1–2 in *b*. *c*, Relations after further magma chamber development. The rate of crystallization is shown as roughly balancing the rate of elevation of the stratified magma so that the top of profile X in *b* and *c* is in contact with magma of about the same degree of compositional evolution. This will result in more or less constant cryptic variations, as is the case above the basal reversal in the Fongen-Hyllingen, Stillwater, Great Dyke and Muskox intrusions. *d*, The crystallization front may be curved, in which case discordant relations between modal and cryptic layering are more likely to develop high on the flanks of the intrusion.

for all the variation in Mg_{opx} . Contamination of the magma near the floor by incorporation of partially molten country rock is possible, but potential partial melts are unlikely to be able to change the $Mg:Fe$ ratios of mafic minerals sufficiently. It is also difficult to envisage how such a mechanism could produce the often very systematic compositional regressions. Other possible explanations for basal reversals include the removal of decreasing amounts of Mg -rich phases in a feeder conduit, or increasing degrees of partial melting of a mantle source. However, these processes do not explain the Stillwater basal reversal¹. Only the first magma batches emplaced into a magma chamber can be envisaged as being evolved as a result of the two latter processes. However, basal reversals appear to occur all along the flanks of the above intrusions, and not only adjacent to the crystalline products of the initial magma influxes. An important feature of most basal reversals is that they show mineral crystallization sequences which are essentially the opposite of those in the overlying layered series (for example, Stillwater: $opx + pl + cpx$, $opx + pl$, opx , ol)⁴. This strongly suggests that they crystallized from similar parental liquids.

While the basal reversals in the Stillwater, Muskox and Great Dyke intrusions are considerably more primitive than at Hyllingen, they all possess features essential to the model presented in Fig. 4. First, repeated magma influx was important during their development^{3,11,12}. Second, both the Great Dyke¹³ and the Muskox² intrusions have inwardly inclined margins, and crystallization along an inclined floor has been suggested for the Stillwater¹² and Muskox¹⁴ intrusions. Furthermore, crystalliza-

tion models involving compositionally stratified magma have been proposed for all three^{3,12,14}.

Therefore, it seems that modern concepts of the fluid dynamic behaviour of magmas can explain the existence of compositional reversals at the base of layered intrusions, although other mechanisms may have some role. A consequence of this process should be that the basal rocks become increasingly evolved up the flanks of the intrusions referred to above. Another consequence should be the presence of discordant relations between modal and cryptic layering on their flanks. This might, however, be difficult to detect since compositional stratification of the magma was fairly restricted during the early stages of development of the magma chambers which are preserved in most of the Stillwater, Great Dyke and Muskox intrusions. Also, the angle of slope of the crystallizing floor, which is probably a curved surface (Fig. 4d), may have been rather small.

The Danish Natural Science Research Council and the Carlsberg Foundation provided financial support to study the Fongen-Hyllingen complex. P. Browning, R. H. Hunter and R. S. J. Sparks, Cambridge University, reviewed the manuscript.

Received 8 May; accepted 21 July 1986.

1. Raedeke, L. D. & McCallum, I. S. *J. Petrol.* **25**, 395–420 (1984).
2. Irvine, T. N. & Smith, C. H. in *Ultramafic and Related Rocks* (ed. Wyllie, P. J.) 38–49 (Wiley, New York, 1967).
3. Wilson, A. H. *J. Petrol.* **23**, 240–292 (1982).
4. Raedeke, L. D. in *Workshop on Magmatic Processes of Early Planetary Crusts: Magma Oceans and Stratiform Layered Intrusions* (eds Walker, D. & McCallum, I. S.) 128–134 (Lunar and Planetary Institute, Tech. Rep. 82–01, Houston, 1981).
5. Wilson, J. R., Esbensen, K. H. & Thy, P. *J. Petrol.* **22**, 584–627 (1981).
6. Wilson, J. R. & Larsen, S. B. *Geol. Mag.* **122**, 97–124 (1985).
7. Klemm, D. D., Ketterer, S., Reichardt, F., Steindl, J. & Weber-Deifenbach, K. *Econ. Geol.* **80**, 1007–1015 (1985).
8. Wilson, J. R. in *The Caledonide Orogen—Scandinavia and Related Areas* (eds Gee, D. G. & Sturt, B. A.) 717–724 (Wiley, New York, 1985).
9. Maaløe, S. *Mineralog. Mag.* **42**, 337–345 (1978).
10. Huppert, H. H., Sparks, R. S. J., Wilson, J. R. & Hallworth, M. A. *Earth planet. Sci. Lett.* (in the press).
11. Irvine, T. N. in *Physics of Magmatic Processes* (ed. Hargraves, R. B.) 325–383 (Princeton University Press, 1980).
12. Irvine, T. N., Keith, D. W. & Todd, S. G. *Econ. Geol.* **78**, 1287–1334 (1983).
13. Worst, B. G., *Bull. S. Rhodesia geol. Surv.* **47**, 1–235 (1960).
14. Irvine, T. N. *Yb. Carnegie Instn Wash.* **80**, 317–324 (1981).
15. Wilson, J. R., Menage, J. F., Pedersen, S. & Engell-Sørensen, O. in *Origins of Igneous Layering* (ed. Parsons, I.) (Reidel, Dordrecht, in the press).

Pleistocene occupation in the south-east Queensland coastal region

Robert Neal* & Errol Stock†

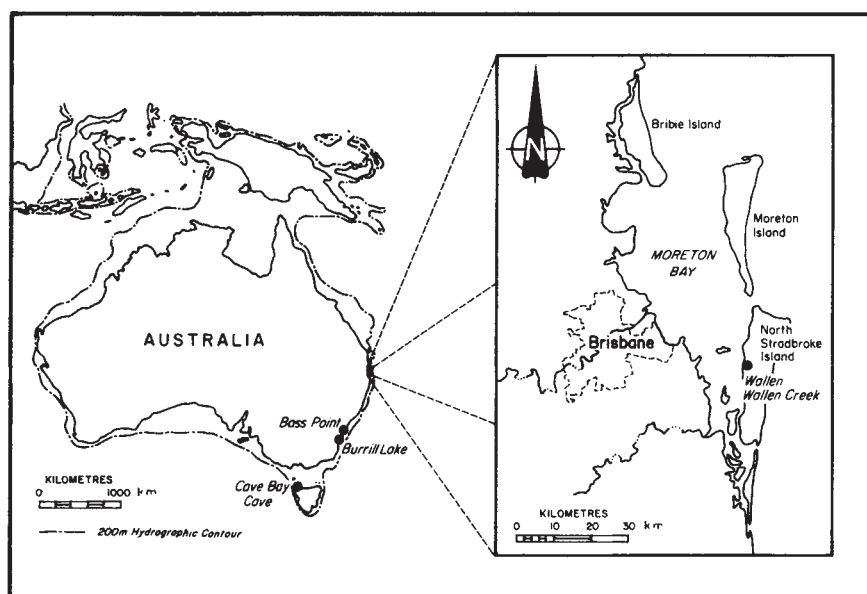
* Department of Anthropology & Sociology, University of Queensland, St Lucia, Q 4067, Australia

† School of Australian Environmental Studies, Griffith University, Nathan, Q 4111, Australia

An open site at Wallen Wallen Creek, North Stradbroke Island has revealed a deep (≥ 2.5 m) stratified archaeological deposit which yielded radiocarbon dates of Pleistocene age. These results are the first evidence of Pleistocene human occupation in south-east Queensland, and the first evidence of Pleistocene occupation of the Australian east coast, north of the Sydney region. The deposit suggests a continuous occupation throughout the Pleistocene, with a dramatic increase in occupation intensity during the late Holocene. Evidence for the continuous occupation of the near-coastal zone during the Pleistocene makes this site unique in the archaeological record of east-coast Australia.

The Wallen Wallen Creek site is located on the west coast of North Stradbroke Island (Fig. 1), on the outside footslope of the left-hand limb of a well-vegetated relict parabolic dune, and adjacent to an extensive fresh-water swamp formed by coastal progradation. The site has a surface elevation of 10.5 m (Australian height datum) and the present low-energy shoreline is ~400 m to the west. Geomorphological analysis has established

Fig. 1 Location of Wallen Wallen Creek.



two soil units formed since the dune was deposited. A giant 6–7-m podsol¹ has formed with a pipey Bh horizon in the upper, well drained part of the ridge, while a humic podsol with a Bh horizon occurs on the poorly drained footslopes, on sand that has moved down slope.

Stratigraphic data² and sea-level curves^{3,4} indicate that the Wallen Wallen Creek site was part of the mainland for most of the past 100,000 years. A sea-level rise of up to 170 m began ~17,000 yr ago, with stabilization near present levels ~6,000–6,500 yr ago^{5,6}. During this period a high dune-field on the mainland, located between a 12–20-km-wide coastal plain to the east and a broad river valley to the west, was transformed to an offshore island, forming part of the eastern boundary of Moreton Bay. After a higher level (+1.35 m) at ~4,000–5,000 yr BP (ref. 7), sea level stabilized and many low-lying islands developed in southern Moreton Bay⁸.

Initial analysis of the Wallen Wallen Creek site has concentrated on an examination of chronology, stratigraphy and the associated archaeological material. The results summarized in Fig. 2 are derived from the analysis of one 1 m × 1 m column (designated WWC-X/Z) and represent one quarter of the material excavated. Eleven radiocarbon dates have been obtained from the WWC-X/Z excavation (Table 1), which confirm and complement those previously obtained from the

original test excavation (designated WWC-M28 and located 5 m to the north).

The excavated column can be interpreted as a continuous occupational sequence containing five well-defined stratigraphic units (Fig. 2). Extending from the present land surface (0 cm) to a depth of 80 cm, Unit 1 comprises a shell midden with vertebrate faunal remains, charcoal and flaked stone artefacts. It gives evidence of a major increase in occupation of the offshore islands of Moreton Bay only during the past few millennia. The inhabitants initially hunted terrestrial and aquatic vertebrate fauna, including dugong (*Dugong dugong*), pademelon (*Thylagale* sp.) and snake (*Python spilates*). This was later replaced by an exclusively coastal economy based on the littoral and marine resources of fish and shellfish, with limited dugong hunting.

Evidence for Pleistocene occupation occurs in Units 2, 3 and 4. Unit 2 (80–170 cm) coincides with the major marine transgression during the late Pleistocene and early Holocene, and contains a few flaked stone artefacts with charcoal and some bone fragments. Between 170 and 225 cm, Unit 3 contains many stone artefacts with charcoal and a stone-tool flaking floor. One of several hearths has been dated (SUA-2341). Unit 4 is a sparse stone scatter between 225 and 305 cm; 225 cm coincides with the upper level of the present water table and 305 cm with the

Table 1 Conventional radiocarbon dates from Wallen Wallen Creek

Sample code	Material	Depth (cm)	Age (yr)	Laboratory Code
Wallen Wallen Creek X/Z excavation				
WWC-ZH-7	Marine shell	22–25	1,070 ± 50	SUA-2461
WWC-ZH-12	Marine shell	33–37	1,230 ± 50	SUA-2465
WWC-ZD-19	Charcoal	58–61	2,120 ± 70	SUA-2462
WWC-ZC-23	Charcoal	69–71	4,290 ± 90	SUA-2466
WWC-ZC-32	Charcoal	97–100	6,950 ± 80	SUA-2343
WWC-ZD-38	Charcoal	114–118	9,760 ± 140	SUA-2467
WWC-X(U+V)–43	Charcoal	138–143	13,040 ± 220	SUA-2463
WWC-ZB-47	Charcoal	158–164	13,650 ± 240	SUA-2342
WWC-X/Z-50	Charcoal	175–179	16,420 ± 260	SUA-2464
WWC-XT-56	Charcoal	201–207	20,560 ± 250	SUA-2341
WWC-XV-75	Sediment	301–310	11,990 ± 70	SUA-2344R
Wallen Wallen Creek test excavation				
WWC-M28A-17+18	Marine shell	49–56	1,520 ± 70	SUA-2236
WWC-M29A-44	Charcoal	136–141	9,810 ± 130	SUA-2251
WWC-M28A-52–54	Charcoal	189–201	17,500 + 900/–800	SUA-2235

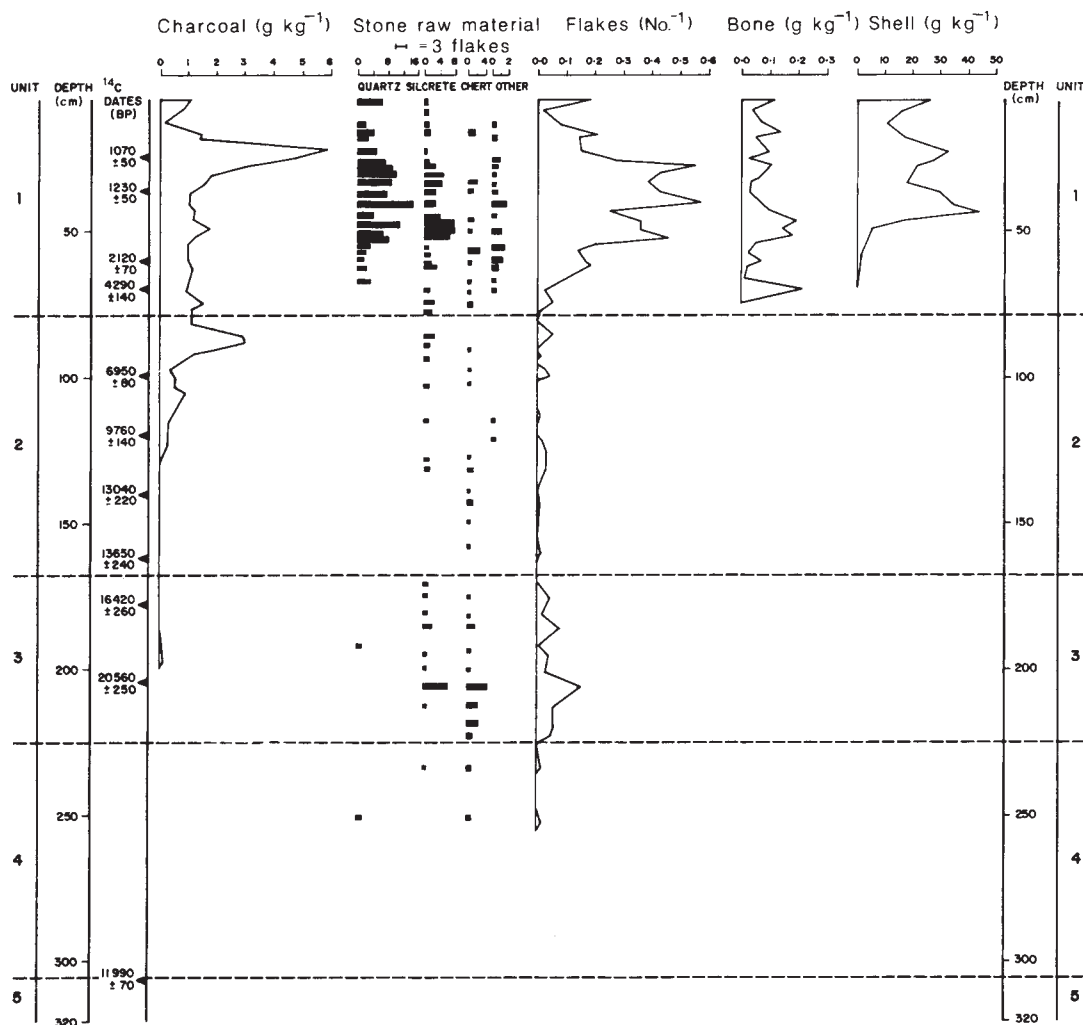


Fig. 2 Summary results from the WWC-X/Z excavation.

top of Unit 5. Stone material within Unit 4 is weathered and shows none of the normal surface morphology of artefacts. The stones are shown to be artefacts, however, by the fact that three stone types do not occur naturally on the island. A humic podsol Bh horizon, designated Unit 5, is a base to the higher archaeological units and appears to be culturally sterile. The radiocarbon date from this unit (SUA-2344R) is anomalous, and is believed to reflect ground-water contamination.

It has long been predicted that the Pleistocene colonists of Australia were coastally oriented⁹, and that the coastal zone was continuously occupied since that time¹⁰⁻¹². Unlike Bass Point¹³, Burrill Lake¹⁴ and Cave Bay Cave¹⁵, however, the Wallen Wallen Creek site confirms that the Australian east coast was continuously occupied during the late Pleistocene. Stone artefacts occur throughout the excavated column, with significant accumulations in Units 1 and 3 (Fig. 2d). In Units 1 and 2 there

is only rare evidence of use-wear and/or retouch, whereas a number of steep-edged scrapers were located in Unit 3 (Fig. 3). The stones in Unit 3 are high-quality chert, silcrete and siliceous petrified wood, which were progressively replaced by quartz, quartzite and silcrete in Unit 2, with the further addition of basalt, rhyolite, quartz conglomerate and sandstone in Unit 1 (Fig. 2c).

The high-quality stone raw materials of the Pleistocene levels, and particularly those from Unit 3, do not occur in the overlying Holocene levels of the site, or elsewhere in the Holocene archaeological deposits surrounding Moreton Bay and dated within the past few millennia¹⁰. The disappearance of high-quality stone raw material from the Moreton Bay archaeological sequence suggests significant changes in raw material availability through time, which may reflect environmental factors such as inundation of the raw material sources during the post-glacial

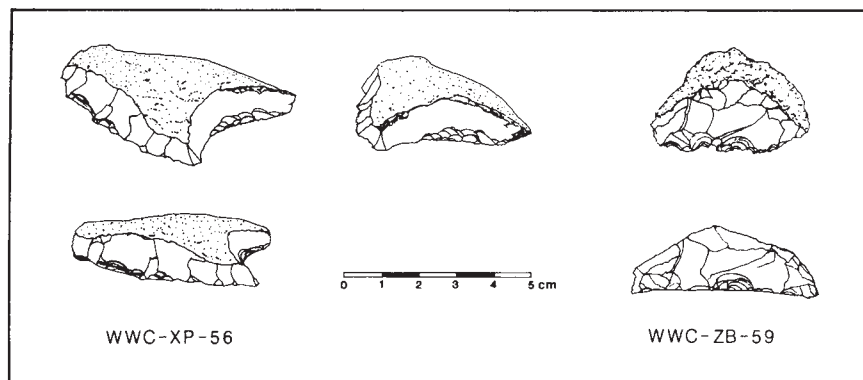


Fig. 3 Artefacts from Unit 3, Wallen Wallen Creek.

marine transgression, or cultural factors such as a decrease in exchange or transportation networks since the Pleistocene.

The location of Wallen Wallen Creek in what must have been a marginal environment throughout the Pleistocene and Holocene, and the limited extent of the archaeological remains, suggests that this site was a temporary transit camp located on an access route between major resource zones located on the coast to the east and the river valley and mountains to the west.

This project was partly supported by the Australian Heritage Commission; the Archaeology Branch, Department of Community Services; and the Department of Anthropology & Sociology, University of Queensland. We thank H. J. Hall, Iain Davidson, Josephine Flood, Sandra Bowdler and Bill Ward for comments on an earlier draft. The figures were drawn by Doug Hobbs and Wayne Steen.

Received 25 March; accepted 1 August 1986.

1. Thompson, G. H. *Nature* **292**, 59-61 (1981).
2. Pickett, J. W., Thompson, C. H., Martin, H. A. & Kelley, R. A. in *Focus on Stradbroke* (eds Coleman, R. J., Covacevich, J. & Davies, P.) 167-177 (Booralong, Brisbane, 1984).
3. Chappell, J. *Search* **14**, 99-101 (1983).
4. Thom, B. G. & Chappell, J. *Search* **6**, 90-93 (1975).
5. Veeh, H. H. & Veevers, J. J. *Nature* **226**, 536-537 (1970).
6. Flood, P. G. in *Monograph Series, Department of Geography, Occasional Papers No. 3* (ed. Hopley, D.) 85-92 (James Cook University of North Queensland, Townsville, 1983).
7. Ward, W. T. & Hacker, J. *Sedimentol. Newslett.* **11**, 54 (1982).
8. Kelley, R. A. & Baker, J. in *Focus on Stradbroke* (eds Coleman, R. J., Covacevich, J. & Davies, P.) 156-166 (Booralong, Brisbane, 1984).
9. Bowdler, S. in *Sunda and Sahul* (eds Allen, J., Golson, J. & Jones, R.) 205-246 (Academic, London, 1977).
10. Hall, J. & Robins, R. in *Queensland Archaeological Research Vol. 1* (ed. Hall, H. J.) 85-94 (Department of Anthropology & Sociology, University of Queensland, St Lucia, 1984).
11. Beaton, J. M. *Archaeol. Phys. Anthropol. Oceania* **20**, 1-20 (1985).
12. White, J. P. & O'Connell, J. F. *A Prehistory of Australia, New Guinea and Sahul* (Academic, Sydney, 1982).
13. Bowdler, S. thesis, Univ. Sydney (1970).
14. Lampert, R. J. *Terra Australis* **1** (Australian National University, Canberra, 1971).
15. Bowdler, S. *Terra Australis* **8** (Australian National University, Canberra, 1984).

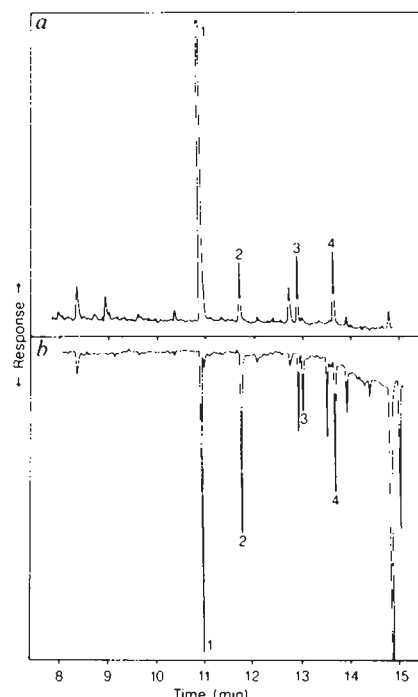


Fig. 1 Gas chromatograms of an airborne collection (a) and a subsequent gland extract (b) from a calling *Y. rorellus* female. Chromatograms displayed as mirror images to emphasize the similarity between the gland contents and the released volatiles. 1, tetradecyl acetate; 2, tetradecanol; 3, hexadecyl acetate; 4, hexadecanol.

Methods. Insects were collected in the field at Ter Aar, near Leiden (location 5 in ref. 6) as pupae or late instar larvae and mailed as pupae to Lund. Males and females were merged and maintained separately on a reversed 16-h light:8-h dark cycle at $\sim 25^\circ\text{C}$ constant temperature. Female moths displayed their pheromone gland in a typical calling posture in the beginning of the photophase about 1 week after emergence. Extracts from the pheromone gland of calling females were prepared in redistilled hexane. Collection of airborne volatiles was carried out according to a procedure modified from ref. 19, and described in detail in ref. 18. The gas chromatograph (Hewlett-Packard model 5880) was equipped with a splitless injector system and a flame ionization detector. A medium polar Supelcowax 10 (crosslinked polyethylene glycol) $30\text{ m} \times 0.25\text{ m.m.}$ inner diameter fused silica column was used. Conditions of chromatography: hydrogen carrier gas flow $40\text{ cm}^3\text{ s}^{-1}$ at 80°C , injector 250°C , splitvalve opened 1 min after injection and temperature maintained at 80°C for 2 min following injection and then programmed at 10°C per min to 230°C .

time as tetradecyl acetate on both a polar capillary column (50% cyano/25% tolyl siloxan) and a medium polar column (Supelcowax 10). The samples consistently contain significant amounts of compounds with retention times identical to those of hexadecyl acetate, tetradecanol and hexadecanol, but no traces of the unsaturated alcohols or acetates found in all other ermine moths can be detected. The gas chromatographic procedure permitted detection of $<50\text{ pg}$, corresponding to $\sim 1\%$ of the major peak. On average, calling females release $5.0 \pm 3.7\text{ ng}$ ($\bar{X} \pm \text{s.d.}$) tetradecyl acetate per 30 min ($n = 12$).

The absence of unsaturated pheromone components was corroborated by analysis of fatty acyl moieties in chloroform-methanol (2:1, v/v) gland extracts after conversion to their methyl esters by base methanolysis⁹. Acetates and alcohols of moth pheromones are biosynthesized from palmitic acid by desaturation and chain shortening¹⁰. The presence of unusual fatty acids in pheromone gland extracts indicates precursors of trace components in the pheromone¹¹, as the acids are often

Evolution of the ermine moth pheromone tetradecyl acetate

C. Löfstedt*, W. M. Herrebout† & J.-W. Du*‡

* Department of Animal Ecology, Lund University, Ecology Building, S-223 62 Lund, Sweden

† Division of Systematics and Evolutionary Biology, University of Leiden, 2300 RA Leiden, The Netherlands

The nine sympatric forms of small ermine moths of the genus *Yponomeuta* (Lepidoptera; Yponomeutidae) in the west palaearctic region show various degrees of differentiation^{1,2}, including among other characteristics, differences in their sexual pheromones³. As is the case for many other moths⁴, the ermine moths so far analysed use delta 11-unsaturated acetates ((Z)-11-tetradecenyl, (E)-11-tetradecenyl and (Z)-11-hexadecenyl acetate) as primary pheromone components. Here for the first time in Lepidoptera we report that a saturated acetate, tetradecyl acetate, is the primary pheromone component⁵ of *Y. rorellus* (Hübner). *Y. rorellus* is almost monomorphic with respect to isoenzyme variation⁶ and has fewer chromosomes than its relatives^{7,8}. We suggest that it has evolved through loss of unsaturated pheromone components in a 'genetic revolution' at a population bottleneck, and by the founders of the new species filling an empty communication niche, separated from that of the ancestral species.

We analysed female pheromone gland extracts in hexane and airborne collections from calling *Y. rorellus* (Fig. 1). The major peak identified by gas chromatography has the same retention

‡ Present address: Shanghai Institute of Entomology, Academia Sinica, 225 Chongqing Road (S), Shanghai, China.

Table 1 Response of male *Y. rorellus* in a flight tunnel to females and blends of synthetic chemicals identified from female glands

Stimulus (μg)	No. tested	Response (%)					
		Wf	Tf	Or	Hw	L	C
Females	8	0	50	25	25	12	12
14:OAc (100)	16	12	81	50	50	31	31
+16:OAc (70)	17	58	94	29	29	29	18
+16:OAc (70) + 14:OH (1) + 16:OH (0.7)	19	40	75	45	35	30	0
16:OAc (70)	8	0	0	0	0	0	0

The flight tunnel was 0.9 m wide $\times$ 0.9 m high $\times$ 2.5 m long. Males were tested with synthetic compounds added to rubber septa (Thomas) in 100 μl solvent. The baits were prepared to release the compounds in the same ratio as calling females. Thus, as half-lives for compounds of different chain lengths differ, we, for instance, used more 16:OAc relative to 14:OAc on the rubber septa than found in the females to obtain similar release rates. The actual ratios released were confirmed by collection of airborne volatiles from synthetic dispensers, as described in ref. 18. The tests were performed at 19 °C and 200 lux, 0–2 h into the photophase of a photoperiod of 16 h light and 8 h darkness. The males were scored for the following behaviours: wing fanning (Wf), take flight (Tf), stationary orientation in the odour plume (Or), flying halfway to the odour source (Hw), landing at the source (L) and copulatory attempts (C). Tetradecyl and hexadecyl acetate were purchased from Sigma and alcohols from Fluka, Buchs. The overall chemical purity of the synthetic samples was >98% and no unsaturated analogues were present according to gas chromatography analysis.

present in larger amounts than the pheromone components. Whereas other ermine moths contain large amounts of delta 11-unsaturated 14- and 16-carbon fatty acids (C.L., W.M.H. & S. B. J. Menken, in preparation), no traces of these compounds can be detected in *Y. rorellus*. Thus we conclude that no delta 11-desaturation takes place in the pheromone gland of this species.

Males tested in a flight tunnel are attracted to tetradecyl acetate alone and combined in natural ratios with hexadecyl acetate, tetradecanol and hexadecanol. These lures elicit upwind flight and landing at the odour source in ~30% of the males tested (Table 1). This level of response is similar to that obtained with mixtures of unsaturated pheromone components in other ermine moths (C.L. and W.M.H., in preparation). Field tests were conducted at Ter Aar, The Netherlands. Sticky delta traps were baited with test chemicals plus BHA (antioxidant) in 100 μl hexane and the traps were dispensed in *Salix* trees, the host plant of *Y. rorellus*. Total trap catches ($n=8$) with tetradecyl and hexadecyl acetate (100+70 μg), tetradecyl acetate (100) and hexadecyl acetate (70) were 73, 96 and 4 males, respectively. The catches with the mixture and with tetradecyl acetate alone were not significantly different according to a Wilcoxon matched-pairs signed-ranks test ($P \gg 0.05$).

Y. rorellus is the only ermine moth in which tetradecyl acetate is behaviourally active without addition of unsaturated acetates. In all others investigated (Fig. 2), subtraction of (*Z*)-11-tetradecenyl acetate (or (*Z*)-11-hexadecenyl acetate in the case of *Y. mahalebells*) destroys the attractivity of the synthetic pheromone. The pheromone of *Y. rorellus* thus differs qualitatively from those of the other species, in which species specificity is largely achieved by specific *E/Z* isomer ratios. Isoenzyme analysis suggests that *Y. rorellus* is on a monophyletic line together with its close relatives *Y. cagnagellus*, *Y. malinellus*, *Y. padellus* and *Y. mahalebells*, and that other lines including *Y. vigintipunctatus*, *Y. plumbellus* and *Y. irrorellus* all have diverged earlier from this group of species than *Y. rorellus*². It thus seems that the simple pheromone of this species evolved through reduction of a more complicated pattern. The alternative but far less attractive hypothesis would be that the delta 11-desaturation operating in most of the ermine moths (and in moths of most other families) evolved on several independent occasions.

The receptor cells on male moth antennae are generally specialized for the detection of their conspecific female-produced pheromone components¹². All the male ermine moths, including *Y. rorellus*, have abundant receptor cells specialized to (*Z*)- and (*E*)-11-tetradecenyl acetate on their antennae¹³, which supports our interpretation that the *Y. rorellus* pheromone evolved from a more complicated pattern by loss of the desaturation ability.

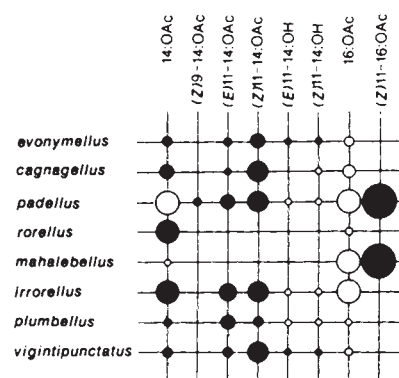


Fig. 2 Pheromone components of eight European small ermine moths. The cross-sectional area of a circle is approximately proportional to the amount of the pheromone component (ng) in an individual female of the respective species. Sizes of circles, starting with the smallest, are graded as follows: 0.4–0.6; 0.7–0.9; 1.0–1.9; 2.0–3.4; 3.5–4.9; 5.0–10; >10 ng. Filled circles, compounds with proven behavioural activity; open circles, compounds present in extracts but with no proven behavioural activity in the respective species. All species except for *Y. rorellus* have delta 11-desaturated compounds as primary pheromone components.

Y. rorellus is almost completely monomorphic at some 75 enzyme loci⁶. The mean proportion of heterozygous loci per individual is 0.003 in *Y. rorellus*, compared with 0.061–0.136 in the other ermine moths². Such a low level of heterozygosity is expected for a species which has been through a bottleneck. Bottlenecks may be important in speciation¹⁴, but the idea of speciation by genetic revolution and the passage through a bottleneck has recently been criticized¹⁵. Theoretical models show that under many assumptions, the probability for the transition of a founder population to a new selective equilibrium, reproductively isolated from the ancestral population, is low. We do not claim that speciation via genetic revolutions and founder effects is the rule or even common, but our results on *Y. rorellus* provide new empirical support for the possibility of such a process. In the case of *Y. rorellus* the aberrant number of chromosome pairs (29 compared with 31 in other species^{7,8}) fits the suggestion of a genetic revolution and it might be that individuals which lost the unsaturated pheromone components entered an empty communication niche to form a new species. This is in principle consistent with the theory of Kaneshiro based on studies of Hawaiian *Drosophila*, which suggests that derived species have a simpler species-recognition mechanism than do ancestral ones^{16,17}. In ermine moths, females of the

founder population should no longer be attractive to the ancestral males as they have lost the primary (unsaturated) pheromone components of the ancestral species. The derived males could still be attracted to the pheromone of the ancestral females, but the differential number of chromosomes could have contributed to reduced fitness of hybrids and as a consequence, the evolution of the asymmetrical reproductive isolation into complete pre-mating reproductive isolation.

We thank H. J. Dijkerman, A. Lindegren, K. Persson and E. Wijk for technical assistance; and B.-O. Bengtsson, P. Brinck and S. B. J. Menken for reviewing the paper. Supported by grants from the Swedish Natural Science Research Council (NFR) and the Netherlands Organisation for the Advancement of Pure Research (ZWO-BION).

Received 2 June; accepted 18 August 1986.

1. Wiebes, J. T. *Neth. J. Zool.* **26**, 440 (1976).
2. Menken, S. B. J. *Z. zool. Syst. Evol. Forsch.* **20**, 131-143 (1982).
3. Löfstedt, C. & Van Der Pers, J. N. C. *J. chem. Ecol.* **11**, 649-666 (1985).
4. Tamaki, Y. in *Comprehensive Insect Physiology, Biochemistry and Pharmacology* (eds Kerkut, G. A. & Gilbert, L. I.) 145-191 (Pergamon, New York, 1985).
5. Roelofs, W. L. & Cardé, R. T. A. *Rev. Entomol.* **22**, 377-405 (1977).
6. Menken, S. B. J. *Evolution* (in the press).
7. Thorpe, W. H. J. *Linn. Soc. Zool.* **36**, 621-634 (1929).
8. Gershenson, Z. S. *Tsitol. Genet.* **1**, 79-80 (1967).
9. Litchfield, C. *Analysis of Triglycerides* (Academic, New York, 1972).
10. Roelofs, W. L. & Bjostad, L. B. *Bioorg. Chem.* **12**, 279-298 (1984).
11. Bjostad, L. B. *et al. J. chem. Ecol.* **10**, 1309-1323 (1984).
12. Priesner, E. in *Chemical Ecology* (ed. Rittner, F. J.) 57-71 (Elsevier, Amsterdam, 1978).
13. Van Der Pers, J. N. C. *Entomol. exp. Appl.* **31**, 255-264 (1982).
14. Mayr, E. *Animal Species and Evolution* (Belknap, Cambridge, Massachusetts, 1963).
15. Barton, N. H. & Charlesworth, B. A. *Rev. Ecol. Syst.* **15**, 133-164 (1984).
16. Ahearn, J. N. *Experientia* **36**, 63-64 (1980).
17. Kaneshiro, K. *Evolution* **34**, 437-444 (1980).
18. Du, J.-W., Löfstedt, C. & Löfqvist, J. *J. chem. Ecol.* (in the press).
19. Baker, T. C. *et al. J. chem. Ecol.* **7**, 961-968 (1981).

Polymorphism of the long-wavelength cone in normal human colour vision

Jay Neitz & Gerald H. Jacobs

Department of Psychology, University of California, Santa Barbara, California 93106, USA

Colour vision is based on the presence of multiple classes of cone each of which contains a different type of photopigment¹. Colour matching tests have long revealed that the normal human has three cone types. Results from these tests have also been used to provide estimates of cone spectral sensitivities². There are significant variations in colour matches made by individuals whose colour vision is classified as normal³⁻⁶. Some of this is due to individual differences in preretinal absorption and photopigment density, but some is also believed to arise because there is variation in the spectral positioning of the cone pigments among those who have normal colour vision. We have used a sensitive colour matching test to examine the magnitude and nature of this individual variation and here report evidence for the existence of two different long-wavelength cone mechanisms in normal humans. The different patterns of colour matches made by male and female subjects indicate these two mechanisms are inherited as an X-chromosome linked trait.

To determine colour matches we used an adaptation of a standard diagnostic test for colour vision, the Rayleigh match⁷. In such matches the subject is required to adjust the proportions of red and green lights until their mixture appears identical to that of a standard yellow light and the mixture and comparison fields usually appear as two halves of a small spot in the central part of the visual field. In this experiment the mixture and comparison lights were presented in Maxwellian view as a large,

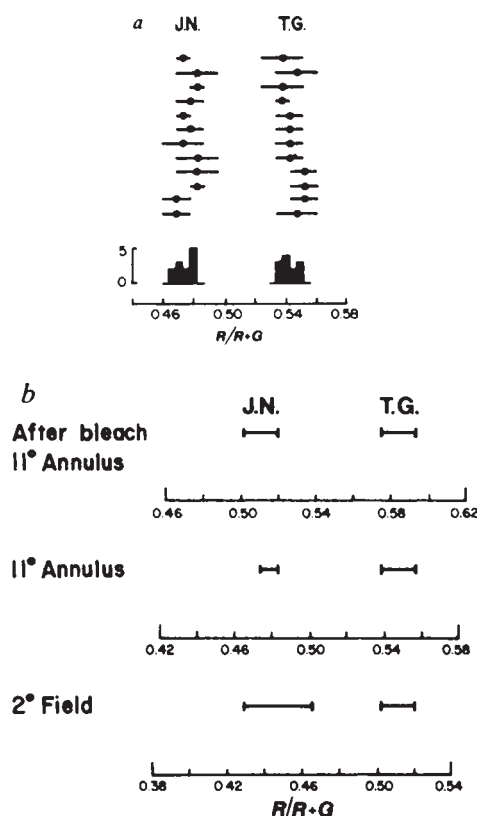


Fig. 1 Rayleigh matches of two male subjects (J.N. and T.G.) with normal colour vision. The stimuli were produced from a three-channel optical system; they appeared in Maxwellian view through an exit pupil having a diameter of 0.7 mm. The mixture lights were 546 nm and 690 nm and the comparison light was 600 nm (each light had a half bandwidth of 10 nm). Each light derived from a separate channel and thus, once the beams were combined, each could be brought to an identical focal point. The mixture and comparison lights were alternated cyclically such that the comparison light appeared for 500 ms, followed by a 100 ms period with no stimulus and then the mixture appeared for 1.9 s. The mixture and comparison lights each produced a retinal illuminance of about 1,000 trolands. Overhead fluorescent lights in the laboratory yielded an additional retinal illuminance of about 500 trolands. *a*, Large field Rayleigh matches of J.N. (left) and T.G. (right) obtained over a period of 12 weeks. The stimuli appeared as an annulus with an outer diameter of 11° and an inner diameter of 3°. Filled circles show the match midpoints; bars indicate the matching ranges. The histograms at the bottom are the distributions of match midpoints. *b*, Rayleigh matches obtained under three stimulus conditions. Bars show the matching ranges of J.N. and T.G. obtained: after a full bleach (top), under the conditions for Fig. 1*a* (middle), and with a 2° foveally-viewed field (bottom).

bright, annular field and were alternated cyclically. Both the large field and the temporal alternation of mixture and comparison lights substantially improve colour discrimination over that obtained in the conventional arrangement⁸. To determine colour matches, the experimenter presented some mixture of red and green light and the subject attempted to match the appearance of the mixture and comparison lights by adjusting the brightness of the comparison light. After adjusting the brightness to satisfaction, the subject was asked to say whether the mixture appeared identical to the comparison light or whether it was redder or greener. This procedure was repeated for a range of red/green mixtures until it was determined which mixtures appeared just reliably redder or greener⁷. The interval between these two points defines the matching range; the centre of this range was taken as the match midpoint.

Table 1 The distribution of Rayleigh matches made by females compared with the distribution expected from matches made by males

Requires:	Expected frequency	Observed frequency
Less red	28	28
Intermediate	50	52
More red	22	20

The prediction is based on the assumption that in individuals with normal colour vision the gene(s) located on the X-chromosome that specify the long wavelength mechanism has (have) two identities. If the frequencies of the two groups of males are a and b , then the predicted frequencies of the three groups of females are given by a^2 , $2ab$ and b^2 . To obtain the observed frequencies, the distribution of female matches in Fig. 1 was divided into three groups at the dips that occur at 0.485 and 0.52.

Application of this matching procedure quickly revealed the presence of individuals who made matches well within the range defined for normal colour vision, but reliably different one from another. That is illustrated in Fig. 1a which shows the match midpoints and matching ranges obtained from each of two male subjects who were repeatedly tested over a period of 12 weeks. The matching ranges for these two subjects are small, but they are similar to those obtained from other subjects tested in this experiment. It can be seen that although the differences in the colour matches made by these two subjects are not large, they are consistently different. That fact is indicated in summary form by the histogram of match midpoints at the bottom of Fig. 1a, which also provides an indication of the small error of measurement in this test situation.

In order to conclude that the differences between the two subjects of Fig. 1a reflect variation in the spectral positioning of their cone pigments, three alternative explanations must be ruled out. (1) The difference is not due to differences in ocular transmissivity. This possibility seemed unlikely because over the wavelengths examined there is little differential absorption by pigments of the lens and macula⁹. To test it directly we measured scotopic spectral sensitivity in these two individuals, and then used these results to derive ocular absorption using the method described by Norren and Vos¹⁰. The resulting difference in ocular absorption so measured was less than 1% of what would be required to produce the difference in colour matches of the two subjects. (2) The colour matching difference is not due to differences in photopigment density of the two subjects. To test this possibility we measured matches for these same two subjects after they had been exposed to a white light sufficiently intense to produce a full bleach of their photopigments¹¹. The resulting matches are shown in Fig. 1b. If the difference in the colour matches made by these subjects were due to differences in the density of their photopigments, their matches should become more similar following the bleach. Although, as expected³, the matches change following the bleach, the matches made by these two subjects do not become more similar. (3) The difference in colour matching of these two subjects is not due to some differences that arise from the use of the large test field. For instance, it has been suggested that photoreceptor orientation has a role in determining the effective optical density of the photopigments and that such effects might be pronounced for large test fields like those used in this experiment^{12,13}. To examine this, colour matches were determined using the same test procedure, but for foveally-viewed stimuli subtending only 2°. The result (Fig. 1b) is that, again, the two subjects continue to show clear differences in their colour matches. It may be concluded that the difference in colour matching of these two subjects is due to a variation in the spectral positioning of their cone pigments.

Colour matches were obtained from a sample of 200 subjects having normal colour vision. Variations in colour matching are

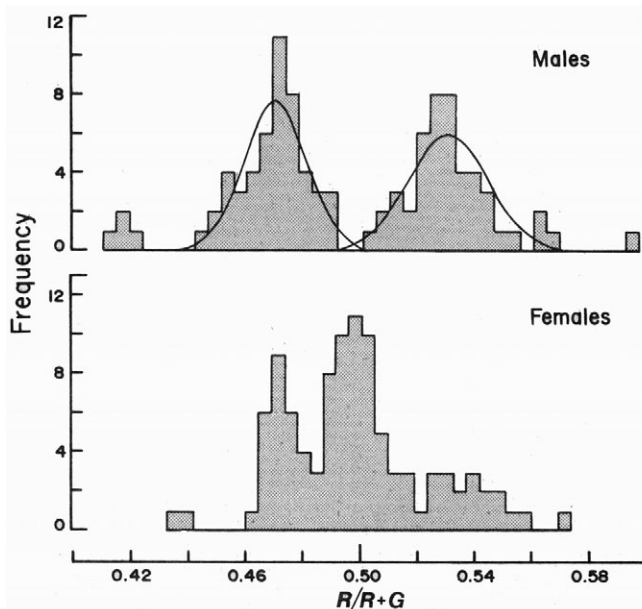


Fig. 2 Distribution of Rayleigh match midpoints for individuals with normal colour vision. The stimulus conditions were identical to those described for Fig. 1a. The R and G values used to compute $R/R+G$ are the energies of the primaries multiplied by constants so that the average match midpoint across all subjects is equal to 0.50. The actual energy values at $R/R+G=0.50$ are $3.21 \mu\text{W}$ for the red light and $1.79 \times 10^{-2} \mu\text{W}$ for the green light. *Top*, the distribution of midpoints obtained for 100 males. The curves outline the normal distributions expected given the mean and standard deviations for each group. The actual energy values determined for the mean match equation for the two groups are as follows: group to the left $R=3.38$, $G=1.65 \times 10^{-2}$, $Y=6.56 \times 10^{-2}$; group to the right $R=3.06$, $G=1.91 \times 10^{-2}$, $Y=6.81 \times 10^{-2}$ (all values in μW). *Bottom*, the distribution of midpoints obtained for 100 females.

associated with differences in ocular pigmentation, the latter covarying with both age and degree of skin pigmentation⁹. To minimize the effects of this source of variability the subjects tested were all young (mean age = 20.6 years) Caucasians. An attempt was made to exclude heterozygous carriers of colour anomaly by rejecting from the sample any females who reported that their families contained colour defective individuals¹⁴.

Frequency histograms of match midpoints are shown separately for male and female subjects in Fig. 2. The distribution of matches gathered from the males is very clearly bimodal; the two subjects of Fig. 1 are individual representatives of these two groups. Assuming that, like the subjects of Fig. 1, the matches made by the subjects of Fig. 2 are also based on variation in the spectral positioning of cones, the modal nature of this distribution indicates that the variation in the underlying photopigments must be discrete. From quantal values of the three components in the mean match equation for each of these two groups, the spectral position of peak absorption (λ_{max}) for two putative cone mechanisms underlying this behaviour can be computed by solving simultaneously the colour matching equations and the polynomial expressions for wavelength dependent nomograms. These computed spectral positions depend on several assumptions. The assumptions made were: first, that these mechanisms have a shape defined by the wavelength-dependent visual pigment nomogram appropriate for A_1 -based pigments having a spectral peak between 530 and 610 nm (ref. 15), second, the extreme long wavelength slopes of the two mechanism follow, respectively, the slopes of the middle and long wavelength cone sensitivity curves measured by Nunn *et al.*¹⁶, third, the photopigments have a specific absorbance of

$0.015 \mu\text{m}^{-1}$ (1) and the photoreceptors in the parafoveal region from 3° to 11° have an average length of $20 \mu\text{m}$ (ref. 17), giving them an axial absorbance at their λ_{max} of 0.3, and fourth, that corrections for lens absorption given in the literature^{18,19} for young humans are appropriate for the present sample. Given these assumptions, for the group of males requiring slightly more red light to complete the match (at the right in Fig. 2) this computation yields spectral peaks of 530 nm for the middle wavelength cone mechanism and 556 nm for the long wavelength cone mechanism. For males requiring slightly less red light the computation yields values of 530 nm and 559 nm for the two mechanisms. There is thus variability in the estimated location of the long wavelength cone mechanism for these two groups, but no corresponding variability in the location of the middle wavelength cone mechanism.

The distribution of matches obtained from females (Fig. 2, bottom) is significantly different from that of the males ($\chi^2 = 77.2$, d.f. = 7, $P < 0.001$). The match most frequently made by female subjects occurs where no male matches. This difference between male and female matches suggests that the variation is X-chromosome linked. The idea is that among normals the gene(s) specifying the spectrum of the long-wavelength cone mechanism has (have) two identities. Males receive either of these and accordingly fall into two groups. Homozygous females would also fall into either of these same two distributions. Assuming the occurrence of X-chromosome inactivation²⁰, females who are heterozygous for this trait would have both types of long wavelength cone mechanism. They thus might be expected to make intermediate colour matches. If this idea is correct, then the distribution of matches made by females is predictable from the distribution of male matches. Table 1 shows this comparison. The distribution of female matches is in close accord with the prediction.

The significant variation in colour matching we find among individuals with normal colour vision is consistent with earlier reports³⁻⁶. The present experiment, however, reveals an unexpected distribution for this variation that leads us to hypothesize that the variation is produced by an X-chromosome linked polymorphism of the long-wavelength sensitive mechanism. Although the trichromacy of normal human colour vision has been established beyond question, the specification of the receptor mechanisms on which the trichromacy is based is not yet complete. This experiment suggests that the long-wavelength cone mechanism assumes at least three distinct identities in different normal human trichromats.

This work was supported by NIH grant EY-02052 and a General Research Grant from the University of California. We thank T. Geist for help in testing the subjects, and J. Fenwick, C. Lyon, K. Krogh and M. Crognale for technical assistance.

Received 10 March; accepted 5 August 1986.

Cholinergic neuropil of the striatum observes striosomal boundaries

A. M. Graybiel*, R. W. Baughman† & F. Eckenstein†

* Department of Brain and Cognitive Sciences, Whitaker College of Health Sciences, Technology and Management, Massachusetts Institute of Technology, 45 Carleton Street, Cambridge, Massachusetts 02139, USA

† Department of Neurobiology, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02115, USA

Acetylcholine and dopamine are key neurotransmitters in the extrapyramidal motor system, where they are thought to lie in a 'functional balance' brought about by interactions between the terminals of the dopamine-containing nigrostriatal tract and the cholinergic interneurons of the striatum¹⁻³. The precise nature of these interactions is not understood, however, nor is it clear how they influence the functioning of striatal systems containing other neurotransmitters. A new clue to understanding such interplay among transmitter-coded systems in the striatum has come from the finding that many of them, including nigrostriatal afferents, follow a macroscopic ordering in which neural elements are concentrated either in or out of the striatal tissue compartments called striosomes⁴⁻⁸. We here report that the cholinergic neuropil of the striatum is also compartmentalized: fibres expressing immunoreactivity to antibodies raised against choline acetyltransferase (ChAT) are sparse in striosomes and are dense in the extrastriosomal matrix. These findings suggest (1) that the interactions between acetylcholine and other neurotransmitters in the striatum are spatially constrained, (2) that cholinergic modulation of striatal function predominates in the extrastriosomal matrix, and (3) that extrapyramidal pathways originating in the matrix, including transthalamic pathways to the frontal lobes, may in particular reflect this cholinergic influence. Such a differential organization of striatal cholinergic circuitry could help to account for the selective therapeutic efficacy of anticholinergic drugs in the treatment of extrapyramidal disorders.

The cholinergic functions of the striatum are thought to depend on the actions of intrinsic cholinergic interneurons^{5,9-11} which correspond to the rare 'giant' cells scattered among the much more common medium-sized striatal-projection neurons^{5,12,13}. The cholinergic perikarya express high levels of ChAT and acetylcholinesterase (AChE) and their distribution in the striatum has been extensively studied by histochemical and immunohistochemical methods^{7,12,13}. It has been more difficult to analyse the cholinergic neuropil of the striatum and to be sure of its origins. Staining of the neuropil for AChE activity is readily achieved, but not all the AChE in the striatum is contained in the cholinergic interneurons¹⁴. Expression of ChAT-immunoreactivity by striatal fibres should define them as cholinergic, but few antisera raised against ChAT have yet permitted reliable immunolocalization of striatal fibre processes^{12,13}. Recently, intense ChAT-immunostaining of the striatal neuropil was found in the rat¹³ and the ferret¹⁵. Some patchiness of staining was noted in the ferret, but the ChAT-positive neuropil was found to be homogeneous in the rat except for broad regional gradients. Taken together with the lack of distinct clustering of cholinergic cell bodies^{12,14}, this suggested¹⁴ that the cholinergic interneuronal system of the striatum might lack the compartmental ordering characterizing most other systems in this region.

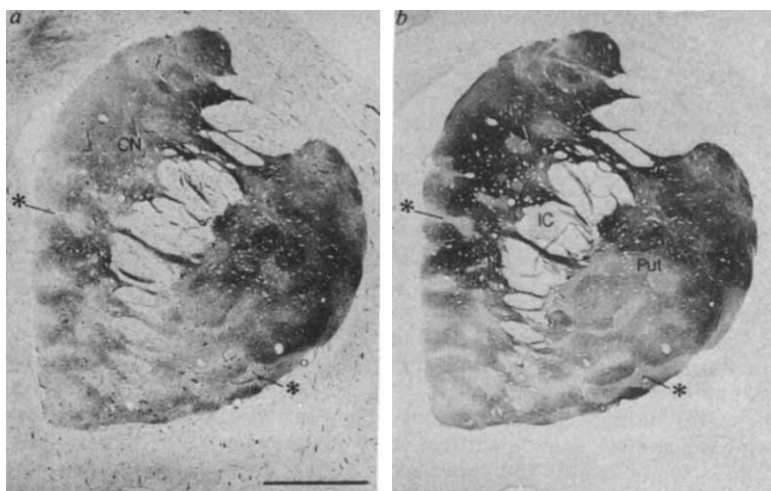
To explore this possibility we used protocols optimizing immunostaining of striatal fibres with antibodies raised in rat against porcine ChAT^{16,17}. We studied the distribution of ChAT-positive neuropil in the striatum of two macaque monkeys and six cats and compared ChAT-immunostaining with AChE histochemical staining in adjoining sections to identify the AChE-poor compartments (striosomes) of the striatum. To test whether

1. Dartnall, H. J. A., Bowmaker, J. K. & Mollon, J. D. *Proc. R. Soc. B* **220**, 115-130 (1983).
2. Boynton, R. M. *Human Color Vision* (Holt, Rinehart and Winston, New York, 1979).
3. Alpern, M. J. *Physiol. Lond.* **288**, 85-105 (1979).
4. Smith, V. C., Pokorny, J. & Starr, S. J. *Vision Res.* **16**, 1087-1094 (1976).
5. Eisner, A. E. & MacLeod, D. I. A. *J. Opt. Soc. Am.* **71**, 705-718 (1981).
6. MacLeod, D. I. A. & Webster, M. A. in *Color Vision Physiology and Psychophysics* (eds Mollon, J. D. & Sharpe, L. T.) 81-92 (Academic, London, 1983).
7. Linksz, A. *An Essay on Color Vision and Clinical Color Vision Testing* (Grune & Stratton, New York, 1964).
8. Nagy, A. L. *J. opt. Soc. Am.* **72**, 571-577 (1982).
9. Wyszecki, G. & Stiles, W. S. *Color Science: Concepts and Methods, Quantitative Data and Formulas* (Wiley, New York, 1967).
10. Norren, D. V. & Vos, J. J. *Vision Res.* **14**, 1237-1244 (1974).
11. Rushton, W. A. H. *J. Physiol., Lond.* **168**, 345-359 (1963).
12. Crawford, B. H. in *Handbook of Sensory Physiology*, Vol. VII/4, *Visual Psychophysics*, (eds Jameson, D. & Hurvich, L. M.) 470-483 (Springer, Berlin, 1972).
13. Nagy, A. L., Purl, K. F. & Houston, J. S. *Vision Res.* **25**, 661-669 (1985).
14. Kalmus, H. *Diagnosis and Genetics of Defective Colour Vision* (Pergamon, Oxford, 1965).
15. Dawis, S. M. *Vision Res.* **21**, 1427-1430 (1981).
16. Nunn, B. J., Schnapf, J. L. & Baylor, D. A. *Nature* **309**, 264-266 (1984).
17. Polyak, S. L. *The Retina* (Univ. Chicago Press, Chicago, 1941).
18. Mellerio, J. *Vision Res.* **11**, 129-141 (1971).
19. Said, F. S. & Weale, R. A. *Gerontologia* **3**, 213-231 (1959).
20. Lyon, M. F. *Am. J. hum. Genet.* **14**, 135-148 (1962).

Fig. 1 A pair of serially adjoining transverse sections through the caudate nucleus (CN) and putamen (Put) of a rhesus monkey illustrating correspondence of staining patterns seen with ChAT immunohistochemistry (a) and acetylcholinesterase histochemistry (b). Asterisks, examples of striosomes originally defined⁴ by weak acetylcholinesterase staining as in b; each striosome can be seen in a as a zone in which ChAT immunostaining is weaker than in the surrounding extrastriosomal matrix. Broad gradients in ChAT and acetylcholinesterase staining are also similar in that staining is stronger dorsolaterally than ventromedially. IC, internal capsule. Bar, 5 mm.

Methods. For ChAT immunostaining, 40- μ m thick frozen sections were pretreated 3×1 h in 100 mM Tris buffer (TBS, 150 mM NaCl containing 2% bovine serum albumin, 10% rabbit serum and 0.5% Triton X-100). They were then incubated: (1) overnight in 1:2,000 rat anti-serum directed against porcine ChAT¹⁶ in TBS;

(2) for 1 h in freshly prepared rabbit anti-rat IgG (Capell) 1:50 in TBS (Triton X-100 at 0.2%) containing 10% conspecific serum; and (3) for 1 h in freshly prepared rat peroxidase-antiperoxidase (Sternberger-Meyer) 1:50 in TBS (Triton X-100 at 0.2%) containing 10% conspecific serum. Steps (2) and (3) were repeated and sections then incubated in diaminobenzidine (DAB, 0.5 mg ml⁻¹), 0.1% H₂O₂ in NaHPO₄ buffer (PB). Steps were separated by 3 5-min TBS washes except before the DAB reaction when PB was used. AChE staining was carried out by a slightly modified Geneser-Jensen and Blackstad method⁴.



the cholinergic cell groups of the basal forebrain and parabrachial midbrain contribute to the striatal neuropil, we made extensive lesions of these cell groups by injecting the excitotoxin ibotenic acid¹⁸ into each of these cell groups 7 days before perfusion (two cats). Brains were fixed by transcardial perfusion with 4% paraformaldehyde and 15% saturated picric acid in 0.1 M Na phosphate buffer (PB, pH 7.4) and were washed extensively and sunk in 20–30% sucrose in PB before being cut and processed for AChE⁶ or for ChAT by a double-bridge peroxidase-antiperoxidase procedure¹⁷ (Fig. 1).

Both in cat and in monkey, well-defined immunoreactive fibres and varicosities were present in the striatum in addition to large ChAT-positive perikarya. The ChAT-positive cell bodies were not obviously clustered. By contrast, heterogeneity in the ChAT-immunoreactive neuropil was immediately evident. In the caudate nucleus of both species and in the ventral parts of the putamen in the monkey, there were conspicuous 0.2–0.6-mm-wide pockets of weak ChAT-immunostaining that stood out against the otherwise densely stained striatal neuropil around them (Fig. 1). Staining in the nucleus accumbens and olfactory tubercle was also highly patterned but more diversely so; there were relatively large and very small (~100 μ m wide) zones of differentially high or low ChAT immunostaining. The islands of Calleja in the olfactory tubercle were especially densely stained¹⁹.

There was a striking correspondence between variations in ChAT immunostaining and variations in striatal AChE staining (Fig. 1). In the caudate nucleus, most zones of low ChAT immunostaining matched the AChE-poor striosomes in shape and location. Both striosomes and less-sharply defined patterning of ChAT immunostaining in the primate putamen also corresponded to the features visible with AChE staining (Fig. 1), although the inhomogeneities were usually clearer in sections stained for AChE. Matches between the ChAT and AChE patterns are also obvious in the ventral striatum.

In addition to this compartmental patterning, there were marked regional gradients in the intensity of ChAT immunostaining of the striatal neuropil. The dorsolateral putamen and associated lateral part of the caudate nucleus, both sensory-motor regions²⁰, showed the densest staining and the ventromedial part of the caudate-putamen complex was least densely stained. These regional variations in ChAT immunostaining were in general mirrored in the patterns of AChE staining.

The ChAT immunostained neuropil included large-diameter dendrites, thin fibres and varicosities of many sizes (Fig. 2). The number of all types of stained structures was reduced in striosomes relative to the number in the extrastriosomal matrix. Some perikarya and dendrites expressing ChAT immunoreactivity did appear inside striosomes, although most lay in the extrastriosomal matrix.

Neither a massive lesion of the cholinergic cell groups of the basal forebrain nor a near-total lesion of the parabrachial cholinergic cell group of the midbrain resulted in any noticeable change in the ChAT-positive neuropil or its striosomal ordering in the ipsilateral striatum. By contrast, ChAT immunoreactivity in the striatum was lost where the ibotenic acid injection site centred in the basal forebrain encroached on the ventral part of the putamen and caudate nucleus.

These experiments point to three conclusions. First, the extrastriosomal matrix is a principal target of cholinergic modulation in the striatum, whereas the striosomal system receives a weaker innervation by ChAT-positive fibres. Second, the cholinergic innervation of the matrix is itself non-uniform and appears especially dense in the sensory-motor region of the striatum.

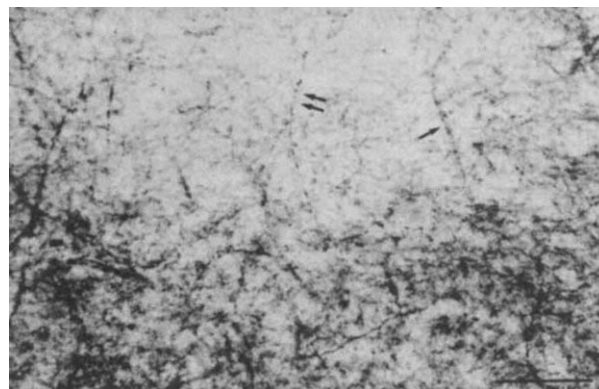


Fig. 2 Change in intensity of ChAT-immunostaining at a border between a striosome (above) and the adjoining extrastriosomal matrix (below). From caudate nucleus of cat. Immunoreactive fibres, dendritic segments and punctata are more numerous in the matrix, although elements of each type do appear in striosomes (single arrow, immunoreactive axon; double arrows, immunoreactive dendrite within striosome). Bar, 50 μ m.

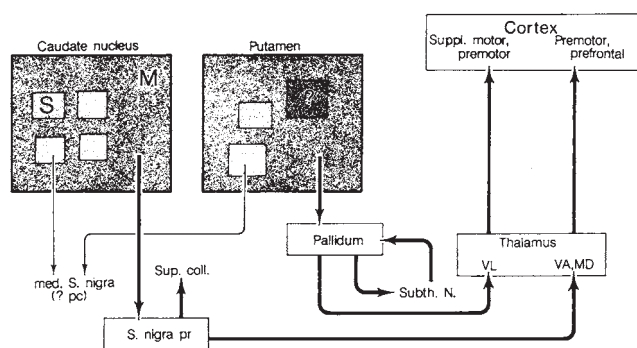


Fig. 3 Summary of extrapyramidal pathways emphasizing with thin arrows and bold arrows, respectively, the different outflow channels emerging from the striosomal (S) and matrix (M) compartments of the caudate nucleus and putamen. Stippling, the finding described here that the density of cholinergic neuropil in the striatum is greater in the extrastriosomal matrix than in striosomes (see text). As a consequence, cholinergic modulation of striatal function should be prominently reflected in pathways of matrix origin. med. S. nigra, medial substantia nigra; pc, pars compacta; S. nigra pr, substantia nigra pars reticulata; Sup. coll. superior colliculus; Subth. N., subthalamic nucleus; VL, ventrolateral nuclei; VA, ventroanterior nuclei; MD, mediodorsal thalamic nuclei; Suppl. motor, supplementary motor cortex. Zone of especially dense stippling in putamen represents zones of apparently intensified ChAT immunostaining (see Fig. 1).

Third, confirming earlier studies^{10,11}, the largest part of this cholinergic neuropil (though perhaps not all of it, see ref. 21) is of intrinsic striatal origin.

Both the local compartmentalization and the regional variation in density of ChAT immunostaining of the striatal neuropil suggest that cholinergic activity in the striatum preferentially affects particular subsets of striatal inputs and outputs (see Fig. 3). On the input side, it is known that cortical afferents from the sensory-motor regions of the neocortex mainly innervate the extrastriosomal matrix^{22,23}, whereas certain inputs related to the limbic system, for example those from the basolateral amygdala²³, mainly innervate striosomes. The cholinergic interneurons of the striatum may therefore have a disproportionate influence on sensory-motor, rather than cognitive-affective, aspects of striatal function.

On the output side of striatal circuitry, there is a sharp division between the striosomal and extrastriosomal systems. Neurones of the striosomal system (at least the dorsal striosomal system identified in the cat) project to the medial part of the substantia nigra, possibly to the dopamine-containing cells of the pars compacta that project back to the striatum^{6,8}. By contrast, neurones of the extrastriosomal matrix project to the pallidum^{8,24} and to the lateral part of the substantia nigra, probably its pars reticulata^{6,8}. The extrastriosomal matrix thus projects to regions that give rise to extrapyramidal input to the thalamus, subthalamus and midbrain tectum and tegmentum.

The fact that the cholinergic neuropil of the striatum is concentrated in the matrix and is weak in striosomes suggests that cholinergic modulation of striatal function would be reflected especially in (1) the subthalamic loop; (2) pathways leading to the midbrain pedunculopontine nucleus and superior colliculus; and (3) pathways leading by way of the pallidum and substantia nigra to the thalamus and then to the premotor and supplementary motor and prefrontal regions of the cerebral cortex (Fig. 3, refs 20, 25).

A clear prediction prompted by this plan of organization is that anticholinergic drugs should strongly affect the thalamic outflow of the corpus striatum-nigral system. This suggestion agrees well with clinical evidence that anticholinergic drugs and stereotactic thalamic lesions have similar effects when applied in the treatment of parkinsonism: both tend to relieve tremor

and rigidity but tend not to affect as much the akinesia associated with the disease²⁶. The regional gradients in ChAT immunostaining in the caudate nucleus and putamen may also be important in accounting for these clinical observations. The dorsolateral putamen and adjoining dorsolateral part of the caudate nucleus receive major inputs from motor and premotor cortex and project preferentially into the pallidal outflow path directed especially to the supplementary motor and premotor cortex^{20,25,27}. The ventral putamen and the large medial and ventral part of the caudate nucleus receive orbitofrontal and prefrontal inputs and project massively to the substantia nigra, whose thalamic outflow pathways are partly directed to more anterior parts of the frontal lobes^{20,25,27-29}.

A second prediction based on these findings is that there is more than one type of cholinergic-dopaminergic balance in the striatum. In particular, it has now been demonstrated in the cat that the dopamine-containing innervations of the striosomes (at least of the dorsal striosomal system) and of the extrastriosomal matrix arise in different subdivisions of the midbrain dopamine-containing cell groups^{8,30}. It has been further reported that binding sites for dopamine receptors follow striosomal ordering, D2 sites being higher in the matrix than in AChE-poor striosomes³¹. Cholinergic-dopaminergic interactions in the striatal matrix may thus represent functioning of the major part of the cholinergic interneuronal system with a subset of its dopamine-containing afferents.

Note that ChAT-positive neuropils in the striosomes, although sparse compared with that in the matrix, is significant and that some cholinergic perikarya and dendrites appear in striosomes. Moreover, one type of muscarinic cholinergic binding site in the striatum (the M1 subtype in cat) is more concentrated inside striosomes than outside them³². The sparse cholinergic innervation of striosomes detectable by immunohistochemistry may therefore reflect a subdivision of the striatal cholinergic system distinct from that of the matrix in both presynaptic and postsynaptic function.

Supported by the Seaver Institute, the Whitaker Health Sciences Fund, the McKnight Foundation, grants NSF BNS 8319547, NIH 2 PO1 NS16367 and NIH 5 RO1 EY02866-08, and the Alzheimer's Disease and Related Disorders Association. We thank Ms D. E. Sahagian and Ms N. Hanley for technical support and Mr H. F. Hall for the photography.

Received 14 May; accepted 21 August 1986.

1. Barbeau, A. *J. med. Ass. Can.* **87**, 802-807 (1962).
2. Douvinois, R. C. *Archs Neurol* **17**, 124-136 (1967).
3. McGeer, P. L., Boudling, J. E., Gibson, W. C. & Foulkes, R. G. *J. Med. Ass. Can.* **177**, 665-670 (1961).
4. Graybiel, A. M. & Ragsdale, C. W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5723-5726 (1978).
5. Graybiel, A. M. & Ragsdale, C. W. in *Chemical Neuroanatomy* (ed. Emson, P. C.) 427-504 (Raven, New York, 1983).
6. Gerfen, C. R. *Nature* **311**, 461-464 (1984).
7. Graybiel, A. M. in *Neuropeptides in Neurologic and Psychiatric Disease* (eds Martin, J. B. & Barchas, J. D.) 135-161 (Raven, New York, 1986).
8. Jimenez-Castellanos, J. & Graybiel, A. M. *Neurosci. Abstr.* **11**, 1249 (1985).
9. Lehmann, J. & Langer, S. Z. *Neuroscience* **10**, 1105-1120 (1983).
10. McGeer, P. L., McGeer, E. G., Fibiger, H. C. & Wickson, V. *Brain Res.* **35**, 308-314 (1971).
11. Woolf, N. J. & Butcher, L. L. *Brain Res. Bull.* **7**, 487-507 (1981).
12. Mesulam, M.-M., Mufson, E. J., Levey, A. I. & Wainer, B. H. *Neuroscience* **12**, 669-686 (1984).
13. Phelps, P. E., Houser, C. R. & Vaughn, J. E. *J. comp. Neurol.* **238**, 286-307 (1985).
14. Fibiger, H. C. *Brain Res. Rev.* **4**, 327-388 (1982).
15. Meininger, C., Rye, D. B. & Wainer, B. H. *Neurosci. Abstr.* **9**, 963 (1983).
16. Eckenstein, F. & Thoenen, H. *EMBO J.* **1**, 363-368 (1982).
17. Eckenstein, F. & Baughman, R. W. *Nature* **309**, 153-155 (1984).
18. Schwarcz, R. *et al. Expl Brain Res.* **37**, 199-216 (1979).
19. Phelps, P. E. & Vaughn, J. E. *Neurosci. Abstr.* **11**, 203 (1985).
20. Alexander, G. E., DeLong, M. R. & Strick, P. L. *A. Rev. Neurosci.* **9**, 357-381 (1986).
21. Saper, C. B. & Loewy, A. D. *Brain Res.* **252**, 367-372 (1982).
22. Donoghue, J. P. & Herkenham, M. *Brain Res.* **365**, 397-403 (1986).
23. Ragsdale, C. W. & Graybiel, A. M. *Neurosci. Abstr.* **10**, 514 (1984).
24. Graybiel, A. M., Ragsdale, C. W. & Moon Edley, S. *Expl Brain Res.* **34**, 189-195 (1979).
25. Graybiel, A. M. *Ciba Fdn Symp.* **107**, 114-149 (1984).
26. Weiner, W. J. & Klawans, H. L. in *Cholinergic-Monoaminergic Interactions in the Brain* (ed. Butcher, L. L.) 335-362 (Academic, New York, 1978).
27. Parent, A., Bouchard, C. & Smith, Y. *Brain Res.* **303**, 385-390 (1984).
28. Ilinsky, I. A., Jouanet, M. L. & Goldman-Rakic, P. S. *J. comp. Neurol.* **236**, 315-330 (1985).
29. Selemon, L. D., Goldman-Rakic, P. S. *J. Neurosci.* **3**, 776-794 (1985).
30. Jimenez-Castellanos, J. & Graybiel, A. M. *Neurosci. Abstr.* **12** (in the press).
31. Marshall, J. F., Joyce, J. N. & Sapp, D. W. *Neurosci. Suppl.* **2**, 207 (1985).
32. Nastuk, M. A. & Graybiel, A. M. *Trends pharmac. Sci. Suppl.* **2**, 92-93 (1986).

Voltage dependence of Na translocation by the Na/K pump

Masakazu Nakao* & David C. Gadsby†

Laboratory of Cardiac Physiology, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

During each complete reaction cycle, the Na/K pump transports three Na ions out across the cell membrane and two K ions in. The resulting net extrusion of positive charge generates outward membrane current^{1,2} but, until now, it was unclear how that net charge movement occurs²⁻⁷. Reasonable possibilities included a single positive charge moving outwards during Na translocation; or a single negative charge moving inwards during K translocation; or either positive or negative charges moving during both translocation steps, but in unequal quantities. Any step that involves net charge movement through the membrane must have voltage-dependent transition rates. Here we report measurements of transient, voltage-dependent, displacement currents generated by the pump when its normal Na/K transport cycle has been interrupted by removal of external K and it is thus constrained to carry out Na/Na exchange⁸. The quantity and voltage sensitivity of the charge moved during these transient currents suggests that Na translocation includes a voltage-dependent transition involving movement of one positive charge across the membrane. This single step can thus fully account for the electrogenic nature of Na/K exchange. The result provides important new insight into the molecular mechanism of active cation transport.

The whole-cell version of the patch-clamp technique⁹ used with wide-tipped pipettes and a device for changing the pipette solution¹⁰ allows control of the compositions of intracellular and extracellular solutions as well as membrane voltage. We have previously used this approach on cells isolated from guinea pig ventricle to measure Na/K pump current, defined as cardiotoxic steroid-sensitive current, after minimizing currents flowing through K, Ca and Na ion channels¹¹. When the pump is strongly activated by intracellular Na at normal extracellular Na and K concentrations, steady-state pump current (200–300 ms after the start of a voltage step) shows marked voltage dependence, but is entirely absent in cells exposed to K-free external solution¹¹ (Fig. 1). The superimposed current records in Fig. 1, *a* and *b*, were elicited by voltage-clamp pulses to potentials between +60 and –100 mV, in a cell equilibrated with 50 mM Na, 10 mM ATP pipette solution, and superfused with K-free, 150 mM Na solution either without (*a*) or with (*b*) 0.5 mM strophanthidin, a cardiotoxic steroid and specific inhibitor of the Na/K pump. The strophanthidin-sensitive currents (Fig. 1*c, d*), obtained by computer subtraction of records in the presence of strophanthidin from those in its absence, confirm that steady-state pump current is zero at all voltages tested.

The strophanthidin-sensitive current associated with a 100-mV depolarization is shown in Fig. 1*c*. The depolarization induces a large transient outward current that decays exponentially; repolarization elicits an inward transient current that decays with a faster exponential time course. Figure 1*d* shows that a hyperpolarizing pulse causes an inward transient pump current, and in that case repolarization is followed by an outward transient.

The absence of steady pump current demonstrates that the electrogenic cycle of Na export and K import by the pump had been halted by withdrawal of extracellular K. Under those conditions, the pump mediates cardiotoxic steroid-sensitive

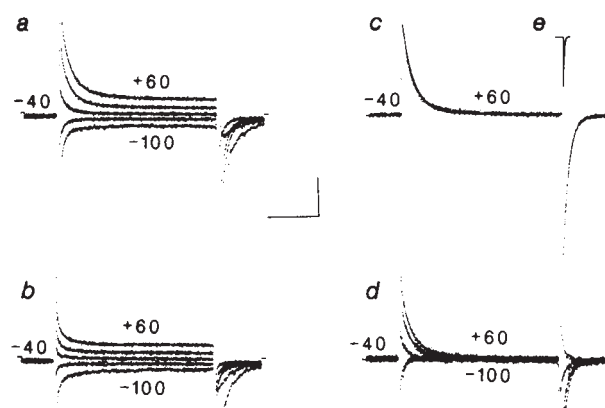


Fig. 1 Strophanthidin-sensitive, transient current in a cardiac cell superfused with K-free fluid. *a, b*, Superimposed records of currents elicited by 100-ms pulses to +60, +20, –20, –60 and –100 mV, from the holding potential, –40 mV, in the absence (*a*) and in the presence (*b*), of 0.5 mM strophanthidin. *c, d*, Strophanthidin-sensitive (pump) currents, obtained by computer subtraction of records in *b* from counterparts in *a*. *c*, Transient pump currents in response to pulse to +60 mV with superimposed, least-squares, exponential fits ($\tau_{on} = 8$ ms, $\tau_{off} = 4$ ms). *d*, Superimposed, transient pump currents for all pulses in *a* and *b*. Inset: *e*, capacity current ($\tau = 0.5$ ms) during 10 mV hyperpolarization, without series resistance compensation; after compensation (as used for all other records), capacity currents decay fully within 1 ms. Repolarization from potentials ≤ -80 mV elicits large Na^+ tail currents which transiently saturate the recording system and prevent analysis of concomitant strophanthidin-sensitive currents. Total cell capacitance, 156 pF. Horizontal bar, 25 ms; vertical bar: *a–d*, 250 pA; *e*, 2.5 nA.

Methods. Quiescent ventricular cells were isolated from adult guinea pig hearts by collagenase digestion and superfused at 34–36 °C with normal Tyrode's solution^{22,23}. We obtained GΩ seals with wide-tipped ($\sim 5 \mu\text{m}$, resistance $\sim 1 \text{ M}\Omega$) pipettes filled with normal Tyrode's which was exchanged¹⁰, just before rupture of the cell membrane, for pipette solution containing (in mM): 50 Na, 80 Cs, 85 aspartic acid, 5 pyruvic acid, 2 MgCl_2 , 10 Mg-ATP , 5 Tris₂-creatine phosphate, 20 TEACl, 5.5 dextrose, 10 EGTA, 10 HEPES (pH 7.4). After establishing voltage control, the cell was superfused with solution (in mM): 150 NaCl, 2.3 MgCl_2 , 2 BaCl_2 , 0.1 CdCl_2 , 5.5 dextrose, 10 HEPES (pH 7.4). The osmolarity of all solutions was 295–300 mosmol kg^{-1} . Strophanthidin was added from a 0.5 M solution in DMSO; up to 0.5% (by volume) dimethyl sulphoxide was shown to have no effect on membrane currents. Current and voltage signals were low-pass filtered at 2 kHz (6-pole Bessel), digitized (12-bit resolution) on-line at 0.15-ms intervals and stored in a computer for analysis.

Na/Na exchange, which is electroneutral in the steady state^{8,12}. Nevertheless, Fig. 2 provides strong evidence that these transient currents are intimately associated with Na/Na exchange. The left-hand sets of strophanthidin-sensitive currents (Fig. 2*A, a; B, a*) were obtained under control conditions, that is, 50 mM Na and 10 mM ATP inside, and 0 mM K and 150 mM Na outside, the cells. The right-hand sets of records were obtained in the same cells within a few minutes of: (1) a switch to pipette solution from which ATP, phosphocreatine, pyruvate and glucose were all omitted (Fig. 2*A, b*); or (2) exposure to Na-free external solution (containing *N*-methyl-D-glucamine instead; Fig. 2*B, b*); or (3) equilibration with Na-free pipette solution (Na replaced by Cs; Fig. 2*B, c*). Each of these manoeuvres abolishes or greatly diminishes the strophanthidin-sensitive transient currents. In additional, preliminary experiments the transient currents are 90% eliminated by $10 \mu\text{g ml}^{-1}$ oligomycin B¹³ and are progressively diminished by addition of extracellular K, with a similar concentration dependence to that for activation of steady-state Na/K pump current¹⁴. These observations all imply a close relationship between the transient currents and Na/Na exchange.

* Present address: Department of Anesthesiology, Hiroshima University School of Medicine, Hiroshima 734, Japan.

† To whom correspondence should be addressed.

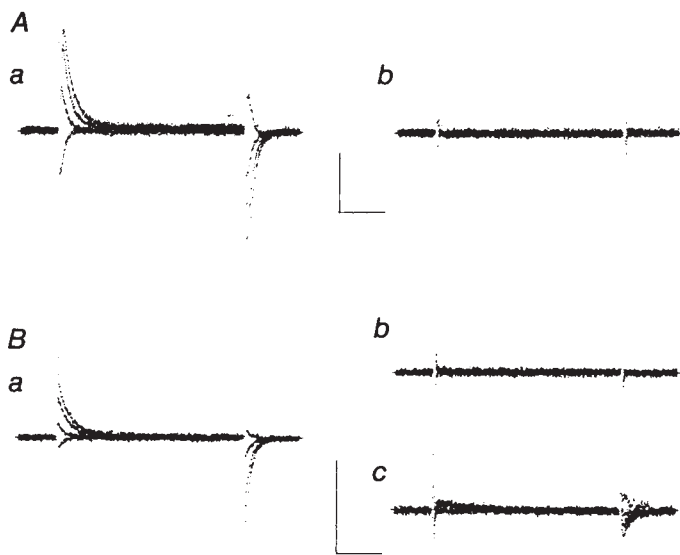


Fig. 2 Transient currents are diminished by removal of internal ATP, internal Na or external Na. *A, B*, Superimposed records of strophanthidin-sensitive currents (obtained as in Fig. 1) elicited by pulses to +60, +20, -20 and -60 mV, from the holding potential, -40 mV. Vertical bar, 250 pA; horizontal bar, 25 ms. *A, a*, Solutions as in Fig. 1; *b*, ~15 min after switching to pipette solution without ATP, creatine phosphate, pyruvate, or dextrose (all replaced by Cs-aspartate); cell capacitance, 194 pF. *B, a*, Solutions as in Fig. 1; *b*, records obtained ~9 min before those in *a*, in the presence of Na-free external solution (150 mM NaCl replaced by 150 mM *N*-methyl-D-glucamine) and 50 mM Na pipette solution (2 mM strophanthidin used in the experiment shown here); *c*, records obtained ~25 min before those in *a* in the presence of 150 mM Na external solution and Na free pipette solution (50 mM Na replaced by 50 mM Cs); cell capacitance, 109 pF. Vertical bar, 250 pA; horizontal bar 25 ms. The small, slow transient current remaining in *c* presumably reflects both the difficulty of maintaining zero intracellular Na concentration against the large inward electrochemical gradient with no means of net Na extrusion, as well as the high affinity of the intracellular pump sites for Na^{24} .

Figure 3 illustrates voltage-dependent characteristics of the transient pump currents. The quantity of charge moved during the transient currents is plotted against pulse potential in Fig. 3*a*, and the rate constant for exponential current decay is shown, on the same voltage scale, in Fig. 3*b*. The rate constant seems to reach a minimum at positive potentials, but increases steeply as the pulse potential is made more negative. (Cooling by 10 °C reduces rate constants at all voltages by a factor of ~3 without altering the quantity of charge moved: data not shown.) The quantity of movable charge (Fig. 3*a*) apparently saturates for large depolarizing or hyperpolarizing pulses. The voltage dependence of the charge moved is reasonably well fitted by a Boltzmann relation (Fig. 3*a*) describing the voltage dependence of the equilibrium population of two states, transitions between which require movement of a single positive charge through the entire membrane field. The total quantity of movable charge in the cell of Fig. 3 is 3.9 pC, or 19 nC μF^{-1} of linear cell capacitance. If a single charge moves per pump, and specific membrane capacitance is assumed to be 1 μFcm^{-2} , then the pump site density is calculated to be 1,200 μm^{-2} , within a factor of two of previous estimates¹⁵.

A simple interpretation of these data (Fig. 3*c*) is that the sites for binding Na ions within the pump molecule provide only two negative charges^{4-7,16}, so that the pump retains a single positive charge when 3 Na ions are bound. The rate of translocation of the Na-loaded sites would then depend on membrane

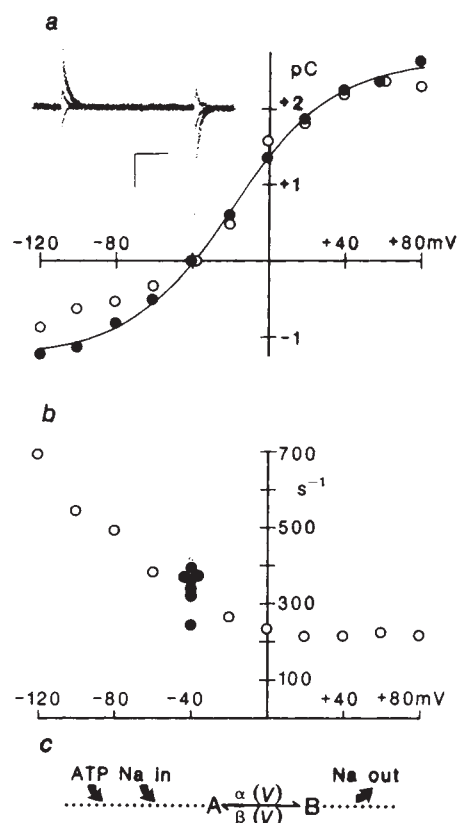


Fig. 3 Voltage dependence of charge movement (*a*) and of decay rate constant (*b*) during transient, strophanthidin-sensitive currents (inset *a*). Intracellular and extracellular solutions as in Fig. 1; holding potential, -40 mV; cell capacitance, 208 pF; temperature, 35 °C. *a*, inset, superimposed sample records of strophanthidin-sensitive currents elicited by pulses to +60, +20, -20 and -60 mV; graph, time integral [$Q(V)$] of current transients initiated by each pulse, plotted against pulse potential (V). Open circles, charge measured directly from the records; filled circles, charge estimated by extrapolation of the exponential current decay. Only 'on' charge movement is plotted; 'off' and 'on' charge movements appear to be the same (for 100-ms pulses to -60 mV < V < +80 mV from a holding potential of -40 mV). The smooth curve, derived from the Boltzmann relation, shows $\Delta Q(V)/\Delta Q_{\max} = 1/[1 + \exp(V' - V)/k]$, where $\Delta Q(V) = Q(V) - Q_{\min}$, $\Delta Q_{\max} = Q_{\max} - Q_{\min}$; V' is the potential at the mid-point; k provides a measure of the equivalent charge. The values used are $Q_{\min} = -1.23$ pC, $Q_{\max} = +2.63$ pC, $V' = -20$ mV, $k = 26.5$ mV. Vertical bar, 250 pA; horizontal bar, 25 ms. *b*, rate constants of exponential fits to transient currents plotted against membrane potential: open circles, rate constants of 'on' transients; filled circles, rate constants of 'off' transients, at -40 mV, following repolarization from -60 mV < V < +80 mV. *c*, simplified kinetic scheme for ATP-dependent Na/Na exchange emphasizing the voltage-sensitive step $A \rightarrow B$, governed by voltage-dependent forward and backward rate constants, $\alpha(V)$ and $\beta(V)$.

voltage. Electrostatic considerations suggest that depolarization should increase the forward, and decrease the backward, rate constant, whereas hyperpolarization should increase the backward, and decrease the forward, rate constant^{3-7,17}, the mid-point potential (V' , Fig. 3*a*) being the voltage at which these rate constants are identical, and the steady-state occupations of states *A* and *B* equal. On suddenly changing the membrane potential, the new equilibrium is approached with an exponential time course whose rate constant (Fig. 3*b*) equals the sum of the forward and backward rate constants at the new voltage. Transient pump currents arise because the rate constants attain their new values instantaneously, whereas the concentrations of *A* and *B* change relatively slowly, resulting in a temporary mismatch between forward and backward charge flux through that step, evident as a transient strophanthidin-sensitive, net

current. This net charge movement declines to zero as the new equilibrium is approached.

We conclude that although Na/Na exchange is an electroneutral process in the steady state at any membrane potential, it includes an essential voltage-dependent step. If this step is rate-limiting for Na/Na exchange, the magnitude of steady-state Na/Na exchange fluxes should vary with membrane potential; such voltage dependence has not yet been detected¹⁸.

Comparison of the forward and backward rate constants suggested by Fig. 3 (near 0 mV) with published data (see, for example, refs 4, 6, 7, 19) raises the possibility that the voltage-dependent transition studied here is the conformational change of the phosphorylated, occluded-Na intermediate: $E_1P \cdot (Na_3) \rightleftharpoons E_2P \cdot Na_3$. Indeed, the nominal absence of ADP (usually required for Na/Na exchange²⁰) in our experiments is consistent with the pump remaining phosphorylated during these voltage-dependent transitions. The same step might limit the rate of the normal forward Na/K pump cycle because, in the same cells, at normal plasma Na and K concentrations steady-state pump current varies steeply with membrane potential¹¹, indicating that a voltage-dependent step in the cycle limits the rate of the cycle. The voltage sensitivity demonstrated in Fig. 3 is sufficient to account fully for forward electrogenic Na/K exchange^{1,2} and probably also for its observed voltage dependence^{4-7,11}. If so, this unexpectedly simple result will facilitate development of kinetic models of the Na/K pump mechanism and, possibly, also of other active transport systems, taking into account recently found similarities²¹.

We thank Dr Paul F. Cranefield for encouragement; Drs Frank Brink, Paul De Weer and Neal Shepherd for critical discussion; and Diane Schenkman for technical assistance. Supported by NIH grant HL-14899 and by an Established Fellowship of the New York Heart Association (D.C.G.).

Received 16 May; accepted 12 August 1986.

1. Thomas, R. C. *Physiol. Rev.* **52**, 563-594 (1972).
2. Glynn, I. M. in *Electrogenic Transport: Fundamental Principles and Physiological Implications* (eds Blaustein, M. P. & Lieberman, M.) 33-48 (Raven, New York, 1984).
3. Hansen, U.-P., Gradmann, D., Sanders, D. & Slayman, C. L. *J. Membrane Biol.* **63**, 165-190 (1981).
4. Chapman, J. B., Johnson, E. A. & Kootsey, J. M. *J. Membrane Biol.* **74**, 139-153 (1983).
5. De Weer, P. in *Electrogenic Transport: Fundamental Principles and Physiological Implications* (eds Blaustein, M. P. & Lieberman, M.) 1-15 (Raven, New York, 1984).
6. Reynolds, J. A., Johnson, E. A. & Tanford, C. *Proc. natn. Acad. Sci. U.S.A.* **68**, 6869-6873 (1985).
7. Luger, P. & Apell, H.-J. *Eur. Biophys. J.* (in the press).
8. Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 159-174 (1967).
9. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflugers Arch. ges. Physiol.* **391**, 85-100 (1981).
10. Soejima, M. & Noma, A. *Pflugers Arch. ges. Physiol.* **400**, 424-431 (1984).
11. Gadsby, D. C., Kimura, J. & Noma, A. *Nature* **315**, 63-65 (1985).
12. Abercrombie, R. & De Weer, P. *Am. J. Physiol.* **235**(1), C63-C68 (1978).
13. Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 217-235 (1967).
14. Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 189-216 (1967).
15. Daut, J. & Rudel, R. *J. Physiol., Lond.* **330**, 243-264 (1982).
16. Karlisch, S. J. D., Raphaeli, A. & Stein, W. D. in *The Sodium Pump* (eds Glynn, I. M. & Ellory, J. C.) 487-499 (Company of Biologists, Cambridge, 1985).
17. Eyring, H., Lumry, R. & Woodbury, J. M. *Rec. Chem. Prog.* **10**, 100-114 (1949).
18. Milanić, M. A. & Hoffman, J. F. *Biophys. J.* **49**, 548a (1986).
19. Glynn, I. M., Hara, Y. & Richards, D. E. *J. Physiol., Lond.* **351**, 531-547 (1984).
20. Glynn, I. M. & Hoffman, J. F. *J. Physiol., Lond.* **218**, 239-256 (1971).
21. Shull, G. E., Schwartz, A. & Lingrel, J. B. *Nature* **316**, 691-695 (1985).
22. Isenberg, G. & Klockner, U. *Pflugers Arch. ges. Physiol.* **395**, 6-18 (1982).
23. Matsuda, H., Noma, A., Kurachi, Y. & Irisawa, H. *Circulation Res.* **51**, 142-151 (1982).
24. Garay, R. P. & Garrahan, P. J. *J. Physiol., Lond.* **231**, 297-325 (1973).

Ammonia orients cell masses and speeds up aggregating cells of slime moulds

J. T. Bonner*, H. B. Suthers* & G. M. Odell†

* Department of Biology, Princeton University, Princeton, New Jersey 08544, USA

† Department of Zoology, University of Washington, Seattle, Washington 98195, USA

We showed some years ago that the rising cell masses of cellular slime moulds repel one another, and that this is achieved by a gas given off from the cell masses¹. It is now clear that this gas is ammonia. We also have evidence that NH₃ tends to speed up the movement of cells in aggregating streams which would account for its ability to repel; more NH₃ on one side of a cell mass would cause a speed-up of the cells on that side, thus making the mass veer to one side.

Cellular slime moulds feed and aggregate in a thin film of water and come together in a cell mass or migrating slug (either with or without a stalk, depending on the species) that moves towards a region favourable for fruiting. Ultimately the migrating slug points upward into the air to form a fruiting body consisting of a terminal spore mass (some species are branched) at the top of a thin, delicate stalk. It is known that during the multicellular stages there may be orientation towards light or towards heat²⁻⁴ and that the slugs will migrate away from more basic regions^{5,6}. In a previous study we were able to show that when the fruiting body (or sorogen, as it is known in its early stages) rises into the air it is guided by a volatile substance that is produced by the sorogen itself¹. As a result, if two sorogens arise close together they lean away from one another; or if one rises near an agar cliff, it will lean away from the cliff. The best evidence that this was due to a gas was obtained by placing a piece of activated charcoal near a sorogen. It would move

Table 1 The effect of different volatile substances on the orientation of the rising sorogens of *D. discoideum*

	Orientation of fruiting bodies		
	Towards hole	Ignoring hole	Away from hole
Water vapour alone	22	12	0
5% CO ₂	14	9	0
1% ethylene	10	4	0
3-5% ethanol	14	8	0
5-15% ethane	17	7	0
0.0036% NH ₃	2	5	29

These experiments were done in a small Petri dish placed in a large desiccator containing the gases listed above, as described in the text and in the legend to Fig. 1.

towards the charcoal, showing that the repellent is adsorbed by the charcoal. At that time we were unable to identify the gas, even though we knew from unpublished work of J. Lonski in our laboratory using gas chromatography, that slime moulds give off CO₂, NH₃, ethane, ethanol, ethylene, acetaldehyde and, no doubt, other volatile substances.

In order to test these chemicals we used small, tightly sealed plastic Petri dishes and drilled a hole 0.35 or 0.4 mm in diameter through the middle of the covers (Fig. 1). An agar block with the primordium of a fruiting body of *Dictyostelium discoideum* (NC-4) was placed below the hole so that the rising sorogen would come out parallel, or slightly away from the undersurface of the lid. This was precisely the result of all control experiments with lids lacking the hole. The small dish was closed (with moist filter paper inside) and placed in a large desiccator which contained the gas to be tested and water to keep the humidity as high as possible, matching the humidity in the small dish. The desiccator was then placed in the dark at 23 °C. Orientation of a sorogen was already evident after an hour, which is approximately the time during which a small, rapidly diffusing molecule can maintain a gradient across the hole under such conditions.

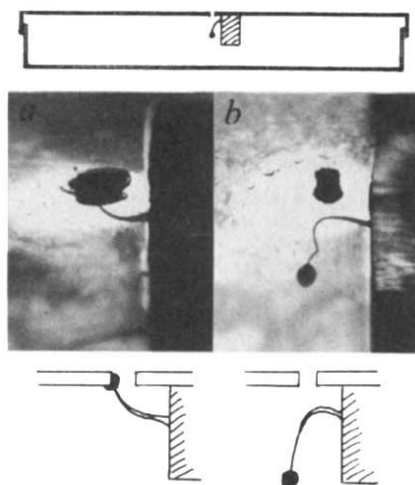


Fig. 1 The top diagram represents a section through a small plastic Petri dish (Falcon No. 1007, 60 × 15 mm) showing the position of the agar block with the slime mould sorogen near the hole in the cover. These small Petri dishes are placed in a large desiccator containing various test gases. The two photographs in the middle show orientation relative to the hole in the top of the Petri dish. In *a* the desiccator contained water vapour only, and in *b* it contained in addition a partial pressure of 0.0027 mm Hg (0.0036%) of NH_3 . The photographs are taken at an angle from above the Petri dish. The source of NH_3 is above the hole. The lower diagrams make clear how the fruiting bodies orient with respect to the holes. (It has been established by Rorke and Rosenthal¹⁶ that gravity plays a negligible role in the orientation of fruiting bodies.) The NH_3 was produced by placing in the base of the desiccator 800 ml of 0.5 M NaOH and 0.002 NH_4Cl . We are indebted to Professor Walter Kauzmann for showing us how to estimate the partial pressure of the NH_3 above this solution. Knowing the volume of the air space in the desiccator (2,190 ml), the temperature (23 °C), and using the Henry's law solubility coefficient derived from the literature¹⁷ for dilute aqueous solutions of NH_3 , it is possible to calculate the partial pressure of NH_3 in the air space. (The same method was used for estimating the partial pressures in the experiments on the effect of NH_3 on the rate of cell movement.)

With water vapour alone as a control, 5% CO_2 , 1% ethylene, 5–15% ethane, or 5% ethanol were allowed to come to equilibrium in the desiccator which contained 800 ml of water in the bottom. In these experiments the sorogen of *D. discoideum* tended to orient towards the hole, presumably because its own repellent could escape through the hole, and it was not affected by the added gases (Table 1, Fig. 1*a*). The result was dramatically different with 0.027 mm Hg (0.0036%) of NH_3 , for the sorogens oriented directly away from the holes (Fig. 1*b*). Clearly NH_3 was a repellent, and it is well known that sorogens give off low concentrations of NH_3 (refs 7–9). The other gases were inactive even at concentrations much above any reasonable estimate of their physiological range.

How does NH_3 orient the sorogen? It has been assumed that the amoebae on one side of the sorogen move more rapidly than on the other, an hypothesis we proceeded to test. Aggregating amoebae of *D. discoideum* on 2% non-nutrient agar in an inverted Petri dish were exposed to mixtures of NaOH and NH_4Cl placed in a small dish inside the cover of the Petri dish. In this way it was possible to test the effect of NH_3 on the rates of movement of amoebae at different stages of aggregation. These rates were measured by using a Panasonic video camera (WV-1850) with time-lapse (AG-6010). We were able to show a significant increase in the rate of movement of the aggregation streams if the air space above the NH_4Cl -NaOH solution had a partial pressure of NH_3 between 0.0025 and 0.0045 mm Hg (Table 2). If the partial pressure exceeded 0.0045 mm Hg the streams broke up and all evidence of cyclic AMP pulses ceased.

Table 2 Effect of NH_3 on the rate of movement of amoebae

	Partial pressure of added NH_3 in mm Hg	Change in speed		No. of cases
		Mean $\pm$ s.d.	Range	
Amoebae in streams	0.0025–0.0045	2.10 times $\pm$ 1.18	1.0 times–6.0 times	23
Isolated amoebae	0.003	1.04 times $\pm$ 0.50	0.5 times–3.0 times	29

Stream speed was measured by placing an acetate sheet over the television screen and marking the outline of the cells, or the streams, at given time intervals (3 min for individual cells; 10 min for streams). It was found that the magnitude of rate increase was variable and did not correlate well with concentration of NH_3 . The measurements were taken either at the very end of a stream, or on the fastest chains of cells within a segment of a stream, care being taken in each case to measure the same region before and after the addition of NH_3 . The change in speed is the speed after treatment divided by the speed before treatment.

Pressures of NH_3 below 0.0025 mm Hg had no apparent effect on the rate of movement of the streams. The speed-up effect was not transient but persisted for an hour or more in the continued presence of NH_3 .

The variability of the response in each experiment is considerable and may stem from the fact that different branches of an aggregate would move at different speeds even without treatment. Furthermore the cells within a stream move at different rates depending on how far away they are from the aggregation centre; cells at the distal end of a stream may move inwards 2.5 times faster or slower than more proximal ones. Also it is clear when one looks across the thickness of a stream that the cells at the edge move more slowly than those running along the centre. There is obviously a slime sheath that surrounds the streams under our conditions, for cells from outside the stream could join only by attaching to the ends of the stream; they were unable to penetrate the side of the stream. All these properties of amoeba locomotion within streams, including the fact that the cells near the stream sides can obtain traction from the slime sheath, are consistent with the model of Odell and Bonner¹⁰ for the locomotion of migrating slugs. The model requires that all of the amoebae secrete some chemical factor that acts to speed up each amoeba's rate of movement.

In some cases it was possible to measure the rate of individual amoebae that were moving towards a stream, and while the effect of NH_3 varied in each experiment, the averaged results showed no net change in their rate of movement before and after the addition of NH_3 (Table 2). Relatively high partial pressures of NH_3 (0.0045 mm Hg) did increase the frequency of the formation of pseudopods and Samuel¹¹ showed that this correlated directly with the speed of movement of slime mould amoebae. From this one can only conclude that the effect of NH_3 on the rate of movement of separate cells (and why it differs from cells in streams) is a matter in need of further investigation.

Since it may easily be shown that, under our conditions, the non-nutrient agar over the NH_3 -generating solution will become considerably more alkaline in a matter of minutes (pH 5.5 to 7.0 in less than 5 min), a series of experiments was run to see if pH changes alone caused the increase in stream speed. Aggregates were allowed to form on squares of dialysis membranes which were initially placed on agar containing 0.02 M Sorensen's buffer adjusted to a pH of approximately 6.0, and after measuring the speed of the cells in a stream the squares were transferred to a similar agar adjusted to a pH of approximately 8.0. The speed of the cells did not change significantly. This was also true if the cellophane squares were shifted from the pH 8.0 agar to the pH 6.0 agar, thus demonstrating that it is the NH_3 and not the external hydrogen ion concentration that is responsible for the changes in the speed reported above.

We did some experiments in the middle of the range of effective NH_3 partial pressures for speeding up aggregation streams (pp NH_3 of 0.0028 mm Hg) and tested for both speed changes and changes in the frequency of the pulses emanating from the aggregation centre before and after treatment, which can easily be counted on the time-lapse video. In three experiments the streams moved from 1.4 to 1.7 times faster, but the average period between pulses either stayed about the same (4.3 to 4.5 min) or increased slightly (4.3 to 6.4 min). This shows that the increase in stream speed cannot be accounted for on the basis of an increase in the frequency of the pulses.

It is known that NH_3 has other effects on the development of cellular slime moulds. Feit⁸ and Lonski⁹ showed that it inhibits aggregation centre formation, and later Schindler and Sussman¹², Thadani *et al.*¹³ and Kay¹⁴ provided evidence that it inhibits the production or the effects of cyclic AMP. Schindler and Sussman¹³ also demonstrated that the presence of NH_3 encourages the migration stage. Finally, Varnum and Soll¹⁵ showed that a constant concentration of cyclic AMP slows the rate of movement of aggregation-competent amoebae, the reverse of the effect of NH_3 reported here. It is evident that NH_3 plays a number of important roles in the control of development of the cellular slime moulds. Furthermore one might be able to explain some of the interesting results of past experiments. For instance, as mentioned earlier, slugs will migrate towards acid regions^{5,6}. This could be because there will be more free NH_3 on their basic side, and thus NH_3 will repel them towards the acid side.

We repeated the main experiments reported here with *Polysphondylium violaceum* in a pilot study and found that while there were differences in detail, *P. violaceum* was repelled by an atmosphere containing 0.03 mm Hg of NH_3 , and that NH_3 also increased the rate of movement of the cells in its aggregation streams. This is consistent with the results of our previous experiment in which it was shown that the repellent gas is not species specific¹.

In conclusion, it is clear that low, physiological concentrations of NH_3 orient rising fruiting bodies and also increase the speed of cells in a stream. This supports the idea that NH_3 is the repellent gas responsible for orientation and that it does so by locally increasing the rate of cell movement in cell masses. Thus NH_3 fits all the criteria needed for the substance postulated by Odell and Bonner¹⁰ that causes cells to move more rapidly in the central axis of a migrating slug.

While this work was in progress we heard from Dr Ira Feit of Franklin and Marshall College that he also has evidence (unpublished) that rising sorogens of *D. discoideum* are repelled by NH_3 , which is an independent confirmation of the result reported here.

We thank E. C. Cox, G. W. Byrne, I. N. Feit and W. J. Kauzmann for helpful comments on earlier drafts. This research was supported by grant DCB-85-2950 from the National Science Foundation and the American Cancer Society to J.T.B. and a grant from the National Science Foundation to G.M.O.

Received 30 May; accepted 29 July 1986.

1. Bonner, J. T. & Dodd, M. R. *Dev Biol.* **5**, 344-361 (1962).
2. Raper, K. B. *J. Elisha Mitchell Sci. Soc.* **56**, 241-282 (1940).
3. Bonner, J. T., Clarke, W. W. Jr, Neely, C. L. Jr & Slifkin, M. K. *J. cell. comp. Physiol.* **36**, 149-158 (1950).
4. Whitaker, B. D. & Poff, K. L. *Expl. cell. Res.* **128**, 87-94 (1980).
5. Raper, K. B. *J. agric. Res.* **58**, 157-198 (1939).
6. Bonner, J. T., Hay, A., John, D. G. & Suthers, H. B. *J. Embryol. exp. Morph.* **87**, 207-213 (1985).
7. Gregg, J. H., Hackney, A. L. & Krivanek, J. O. *Biol. Bull.* **107**, 226-235 (1954).
8. Feit, I. N. thesis, Princeton Univ. (1969).
9. Lonski, J. *Dev Biol.* **51**, 158-165 (1976).
10. Odell, G. M. & Bonner, J. T. *Phil. Trans. R. Soc. Lond.* **B312**, 487-525 (1986).
11. Samuel, E. W. *Dev Biol.* **3**, 317-335 (1961).
12. Schindler, J. & Sussman, M. *J. molec. Biol.* **116**, 161-170 (1977).
13. Thadani, V., Pan, P. & Bonner, J. T. *Expl. Cell Res.* **108**, 75-78 (1977).
14. Kay, R. R. *J. Embryol. exp. Morph.* **52**, 171-182 (1979).
15. Varnum, B. & Soll, D. R. *J. Cell Biol.* **99**, 1151-1155 (1984).
16. Rorke, J. & Rosenthal, G. thesis, Princeton Univ. (1959).
17. Baldi, G. & Specchia, V. *La Chimica e l'Industria* **53**, 929-933 (1971).

Flavones induce expression of nodulation genes in *Rhizobium*

John W. Redmond*, Michael Batley*,
Michael A. Djordjevic, Roger W. Innes*,
Peter L. Kuempel* & Barry G. Rolfe†

Genetics Department, Research School of Biological Sciences,
Australian National University, PO Box 475, Canberra,
ACT 2601, Australia

Successful establishment of nitrogen-fixing root nodules¹ in legume plants requires the presence of several plasmid-borne genes in the *Rhizobium* symbionts²⁻¹². Those genes so far defined (designated *nodABCDEF*) are arranged in three separate operons (*nodABC*, *nodD* and *nodFE*), based on DNA sequence analysis and expression data using *nod* genes fused to *Escherichia coli lac* genes⁴⁻¹⁰. Only one gene, *nodD*, is expressed by *Rhizobium* in culture, and its gene product, together with factors secreted by the plant host, induces expression of genes in *nodABC* and *nodFE*⁷⁻⁹. In addition, *lac* fusions to presently undefined *nod* genes, located in regions II and IV of the 14-kilobase (kb) nodulation fragment of *Rhizobium trifolii*, also respond to the presence of plant exudate⁷. We report here that flavones found in washings of undamaged clover roots induce *nod* gene expression. The flavones are not the same as other plant factors recently identified as inducers of the expression of virulence genes in the plant-pathogenic bacterium *Agrobacterium tumefaciens*^{13,14}, which suggests that there may be a unique set of signals for each type of plant-bacterium interaction. This result has profound implications for the understanding of these interactions and provides strategies for investigating the control of nitrogen fixation and plant disease.

Infection of legumes by *Rhizobium* occurs mainly in the emerging root hair zone¹⁵. To demonstrate that this is the site of release of the substances that induce *nod* gene expression, seedlings were placed on a lawn of a *R. trifolii* strain containing *lac* genes from *E. coli* fused to the *nodA* gene (Fig. 1), or to genes in the *nodFE* operon or to region II or IV. Expression of *nod* genes in the presence of a plant signal led to production of β -galactosidase and development of a blue colour in the presence of 5-bromo-4-chloro-3-indolyl β -D-galactoside (X-gal). Intense blue areas extended 10-20 mm from the root tips, cotyledons and cut ends of dissected seedlings. Colour was not produced in the absence of either the plants or the bacteria, or when bacteria containing an anti-sense *nod::lac* fusion were used (Fig. 1). Within the blue area surrounding the legume roots was a distinct, but variable, clear zone (Fig. 1), which indicates either inhibition by higher concentrations of the plant factor or production of varying amounts of a second, inhibitory, factor which abolishes *nod* gene expression. Seedlings of pea, bean, soybean, alfalfa and clovers all released stimulatory factor(s), but the non-legumes maize, rice and *Parasponia*, did not. The failure of *Parasponia* to produce the factor is surprising because it is nodulated by several *Rhizobium* strains¹⁶. However, the mode of infection of *Parasponia*, characterized by 'crack entry' and a general lack of release of bacteria from ramifying infection threads, is rather different from that observed in most legumes^{16,17}.

The stimulatory substances could be extracted from aqueous washings of seedlings of white clover, *Trifolium repens*, at pH 8.5, but not at pH 11.0, indicating the presence of phenolic groups and the absence of an acidic function. Furthermore, they could be adsorbed to, and eluted from, a column of silicic acid. The stimulatory substances were subsequently isolated from ethanol

* Permanent addresses: School of Chemistry, Macquarie University, North Ryde, New South Wales 2113, Australia (J.W.R., M.B.); Department of MCD-Biology, University of Colorado, Boulder, Colorado 80309, USA (R.W.I., P.L.K.).

† To whom correspondence should be addressed.

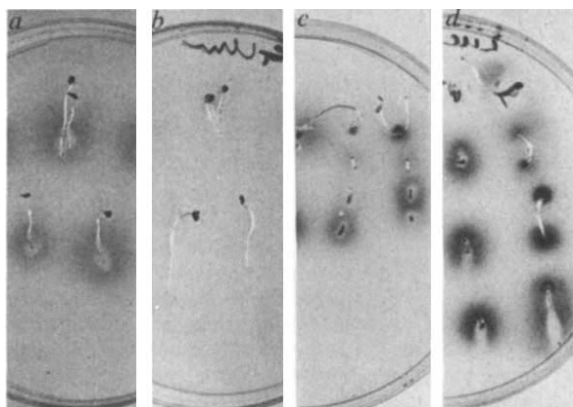


Fig. 1 Bioassay demonstrating the presence of plant-secreted factors which stimulate and inhibit *nod* gene expression in *R. trifolii*. Bacterial lawns of mutants *nod218* (*nodA*::*lac* fusion) and *nod802* (*nodC* anti-sense *lac* fusion) were incorporated into soft agar overlays of nitrogen-free Fahraeus medium. The lawns were spread onto Fahraeus medium containing X-gal (4 mg ml⁻¹). Axenically grown white clover or alfalfa seedlings, or dissected segments of these seedlings, were placed on the bacterial lawns and incubated at 29 °C in the dark for 18 h. The darkened (blue) areas around the clover seedlings (a) indicate that expression of the *nodA*::*lac* fusion occurs only in the vicinity of the roots. A distinct clear area, indicating inhibition of *nod* gene expression, can also be seen. No expression of the *nodC* anti-sense *lac* fusion (b) was observed under the same conditions. Dissected segments of white clover (c) and alfalfa (d) seedlings released stimulatory and inhibitory factors, suggesting that these are transported systemically throughout the plants. Transcriptional fusions to the *R. trifolii* *nod* genes, located on plasmid pRt032, were isolated using the mini-Mu bacteriophage transposon MudI 1734 (ref. 7). The locations of mutants *nod218* and *nod802* are indicated in Fig. 4.

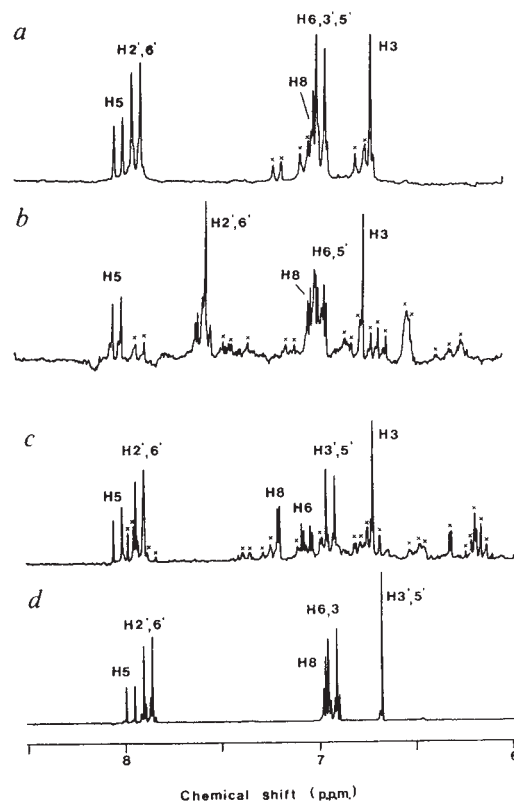


Fig. 2 Proton nuclear magnetic resonance spectra of active fractions from root extracts. a–c, Spectra of the fractions A–C, respectively, dissolved in methanol-d₄. Chemical shifts were measured relative to the residual methanol-d₃ signal at 3.30 p.p.m. The major components are assigned as a, 7,4'-dihydroxyflavone; b, 7,4'-dihydroxy-3'-methoxyflavone (geraldone); c, 4'-hydroxy-7-methoxyflavone; and d, synthetic 7,4'-dihydroxyflavone. The peaks marked with × are due to the presence of (unidentified) impurities.

Methods. The active fractions were isolated from *Trifolium repens* by growing sterilized seeds for 2 days in the dark. The seedlings were washed with distilled water and extracted with 40% ethanol for 4 h, then 2 vol. ethanol were added and the mixture allowed to stand at 4 °C overnight to precipitate proteinaceous material. The supernatant was evaporated to dryness, then taken up in methanol. Polar components were precipitated with 2 vol. chloroform and the supernatant eluted from a Kieselgel 60 (Merck) column with chloroform/methanol (9:1). Active components were isolated by re-chromatography on a column of Kieselgel 60 (Lobar, Merck) using chloroform/methanol (96:4). Fractions (1 ml) were batched on the basis of TLC (Kieselgel Alufolien, Merck) in the same solvent.

extracts of washed seedlings by chromatography on silicic acid. Eluates were combined on the basis of TLC. Three active fractions, labelled A–C in order of increasing mobility, were obtained and the major component in each (~80% as judged by nuclear magnetic resonance spectroscopy) was a substituted flavone (Fig. 2). The major component of fraction A had spectroscopic (NMR, ultraviolet and mass spectra), chromatographic (HPLC, TLC) and fluorescence properties identical to those of an authentic sample of 7,4'-dihydroxyflavone (DHF; synthesized by Professor R. A. Eade, University of New South Wales). The identity of the component was further confirmed by acetylation followed by NMR spectroscopy. The spectra showed the presence of two acetyl groups and the changes in proton chemical shifts were as expected for acetyl substituents¹⁸.

Proton NMR spectra were used to identify the major components in fractions B and C. Spin-spin splitting patterns, confirmed by homonuclear decoupling experiments, showed unambiguously the distribution of substituents on the aromatic rings. (The multiplet pattern for fraction B, however, was complex and simulation of the spectrum was required to verify the assignments.) The chemical shift for each proton could then be used to deduce the nature of the substituents¹⁸. The active components in fractions B and C were thus assigned as geraldone and 4'-hydroxy-7-methoxyflavone, respectively. Sample B had the same chromatographic properties as previously reported for geraldone¹⁹. The natural occurrence of 4'-hydroxy-7-methoxyflavone has not been reported previously, but the NMR spectrum with its singlet at 6.72 p.p.m., characteristic of a flavone²⁰, and the downfield shifts of the A-ring protons compared with those in 7,4'-dihydroxyflavone, leaves little doubt about its chemical identity. The mass spectrum contained peaks of relative molecular mass (*M_r*) 268 and 240 and the higher chromatographic mobility of sample C is consistent with the presence of a single free hydroxyl group. All three samples exhibit the strong

blue fluorescence, modified by exposure to ammonia, that is characteristic of flavones lacking a substituent in the 5 position²⁰.

To confirm that 7,4'-dihydroxyflavone was the active component in fraction A, the biological activity of purified synthetic DHF was compared with that of fraction A (Fig. 3a), using the ultraviolet absorption at 330 nm to estimate the concentration of DHF in fraction A. The expression of *nod* genes in bacteria containing a *nodA*::*lac* fusion (*nod218*)⁷ was detected at a concentration of 7,4'-dihydroxyflavone of 3×10^{-8} M and increased strongly with increasing DHF concentration, reaching a level more than 20 times the background at 6×10^{-6} M (Fig. 3a). Results obtained at higher concentrations were unreliable because of the limited solubility of the compound in water. Induction of *nodA* gene expression was also obtained with fraction B, and to a lesser extent with fraction C of the root extract. The substances were active in the same concentration range as 7,4'-dihydroxyflavone (concentrations were estimated by assuming that the ultraviolet absorbance at 330 nm was the

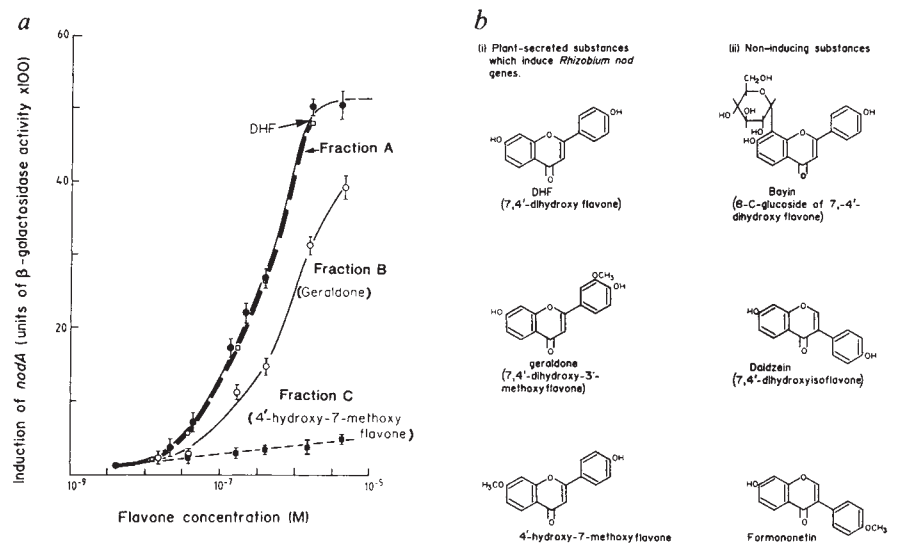


Fig. 3 The structures and *nod* gene-inducing activities of flavones and related compounds. **a**, The levels of induction of β -galactosidase activity in *R. trifolii* cells possessing a *nodA::lac* fusion (*nod218*) after exposure to various concentrations of inducer substances: fractions A ($\square$), B ($\circ$) and C ($\blacksquare$) and authentic 7,4'-dihydroxyflavone ($\bullet$). **b**, Structures of the substances tested.

Methods. Dilutions of each substance, prepared in 95% ethanol, were added to Eppendorf tubes and the solvent carefully removed in a nitrogen stream. Water (1.2 ml) was added and the tubes agitated vigorously before addition of a suspension (0.40 ml, 5×10^8 cells) of early log-phase cells of mutant *nod218*. After 4 h at 29 °C, a sample (0.4 ml) was removed, lysed with chloroform/SDS²⁴ and incubated for 30 min at 42 °C to inactivate the native *Rhizobium* β -galactosidase activity. The induced levels of β -galactosidase activity were then determined using *O*-nitrophenyl β -D-galactoside as substrate²⁴.

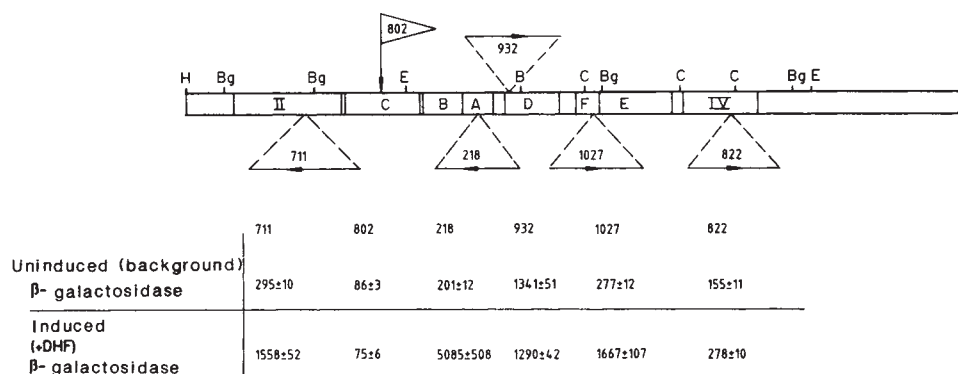


Fig. 4 Locations and orientations of *nod* gene/*lac* fusions in *R. trifolii* and their induction by synthetic 7,4'-dihydroxyflavone. MudI 1734-induced mutants in plasmid pRt032 were isolated as described previously⁷. The fusions of MudI 1734 to *nodA* (*nod218*), *nodD* (*nod932*), *nodF* (*nod1027*), and to regions II (*nod711*) and IV (*nod822*), are indicated. The location of the anti-sense *nodC::lac* fusion (*nod802*) is also shown. Inductions were carried out as described in Fig. 3 legend, with and without sterile solutions of DHF.

same as that for 7,4'-dihydroxyflavone), but the maximal effects were different (Fig. 3a); this may be the result of either solubility limitations or the requirement of a 7-hydroxy function for optimal activity.

The other genes in the nodulation cluster were also induced by plant substances. 7,4'-dihydroxyflavone at 5×10^{-7} M induced transcription of genes of the *nodFE* operon, as well as those in regions II and IV of *R. trifolii* (Fig. 4).

Several other compounds (Fig. 3b) were tested for their ability to induce *nod* gene transcription. No activity was found for formononetin and umbelliferone (which were found in the root extracts) or for daidzein (7,4'-dihydroxyisoflavone), coumestrol and bayin²¹ (which were not). Bayin is the 8-C-glucoside of 7,4'-dihydroxyflavone, and its inactivity suggests fairly specific requirements for the molecular structure of inducer molecules.

The stimulatory factors produced by the other legume seedlings have not yet been characterized, but 7,4'-dihydroxyflavone is known to be present in other legumes²². In clovers DHF exists predominantly as the 7-O-glucoside²³, but analytical chromatography¹⁹ of our root extracts showed no detectable amounts of the glycosides, and acid hydrolysis of the extracts did not release more of the free phenols. The release of stimulatory compounds from the ends of dissected seedlings (Fig. 1) suggests that these compounds are transported throughout the plant, probably via the vascular tissue, but are secreted in appreciable amounts only from tissues in the roots and cotyledons.

Reported natural sources of 7,4'-dihydroxyflavone and geraldone are confined to legumes²¹, which is consistent with our

inability to detect signal substances from non-legume plants. If the compounds are restricted to this family, the ability of *Rhizobium* cells to express nodulation genes in response to particular flavones may represent an adaptation that gives the *Rhizobium* a competitive advantage in the legume rhizosphere. Optimal concentrations of the flavones occur near the emerging root hair zone (Fig. 1), which is the most favourable site for *Rhizobium* infection. If non-leguminous plants could be induced to make similar flavones to those reported here, it might be possible to extend the host range of *Rhizobium*.

We thank Anne Moten and Jan McIver for technical assistance, Professors R. A. Eade and J. Stevens of the University of New South Wales for supplying bayin and synthetic 7,4'-dihydroxyflavone, and Professor Brian Gunning for reading the manuscript. This work was supported in part by an Agrigenetics Sponsored Research programme.

Received 21 February; accepted 8 May 1986.

- Vincent, J. M. in *A Treatise on Dinitrogen Fixation* Vol. 3 (eds Newton, W. E. & Orme-Johnston, W. H.) 277-364 (University Park Press, Baltimore, 1980).
- Djordjevic, M. A., Schofield, P. R. & Rolfe, B. G. *Molec. gen. Genet.* **200**, 463-471 (1985).
- Downie, J. A., Knight, C. D., Johnston, A. W. B. & Rossen, L. *Molec. gen. Genet.* **198**, 255-262 (1985).
- Egelhoff, T. T., Fisher, R. F., Jacobs, T. W. & Long, S. R. *DNA* **4**, 241-248 (1985).
- Torok, I., Kondorosi, E., Strepkowski, T., Postai, J. & Kondorosi, A. *Nucleic Acids Res.* **12**, 9509-9524 (1984).
- Rossen, L., Johnston, A. W. B. & Downie, J. A. *Nucleic Acids Res.* **12**, 9497-9508 (1985).
- Innes, R. W. *et al. Molec. gen. Genet.* **201**, 426-432 (1985).
- Mulligan, J. T. & Long, S. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6609-6613 (1985).

9. Rossen, L., Sherman, C. A., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **4**, 3369-3373 (1985).
10. Rolfe, B. G. *et al* in *Advances in the Molecular Genetics of Bacterial-Plant Interaction* (eds Szalay, A. A. & Legocki, R. P.) 43-48 (Media Services, New York, 1985).
11. Kondorosi, E., Banfalvi, Z. & Kondorosi, A. *Molec. gen. Genet.* **193**, 445-452 (1984).
12. Djordjevic, M. A. *et al* *Pl. molec. Biol.* **4**, 147-160 (1985).
13. Okker, R. J. H. *et al* *Nature* **312**, 564-566 (1984).
14. Stachel, S. E., Messens, E., Van Montagu, M. & Zambryski, P. *Nature* **318**, 624-629 (1985).
15. Bauer, W. D. A. *Rev. Pl. Physiol.* **32**, 407-449 (1981).
16. Trinick, M. J. & Galbraith, J. *New Phytol.* **86**, 17-26 (1980).
17. Lancelle, F. A. & Torrey, J. G. *Protoplasma* **123**, 26-37 (1984).
18. Williams, D. H. & Fleming, I. *Spectroscopic Methods in Organic Chemistry* 135 (McGraw-Hill, London, 1973).
19. Wong, E., Mortimer, P. I. & Geissman, T. A. *Phytochemistry* **4**, 89-95 (1965).
20. Markham, K. R. & Mabry, T. J. in *The Flavonoids* (eds Harborne, J. B., Mabry, T. J. & Mabry, H.) 267-295 (Chapman & Hall, London, 1981).
21. Eade, R. A., Salasoo, I. & Simes, J. J. H. *Aust. J. Chem.* **19**, 1717-1727 (1966).
22. Venkataraman, K. in *The Flavonoids* (eds Harborne, J. B., Mabry, T. J. & Mabry, H.) 267-295 (Chapman & Hall, London, 1981).
23. Wong, E. & Francis, C. M. *Phytochemistry* **7**, 2123-2129 (1968).
24. Miller, J. H. *Experiments in Molecular Genetics*, 352-355 (Cold Spring Harbor Laboratory, New York, 1972).

Transcripts of functionally rearranged gamma genes in primary T cells of adult immunocompetent mice

Barry Jones*, Shelley Mjolsness†, Charles Janeway Jr* & Adrian C. Hayday†

* Department of Pathology and Howard Hughes Medical Institute, and † Department of Biology, Yale University, New Haven, Connecticut, USA

The T-cell specific, rearranging γ -chain genes bear striking resemblance to T-cell receptor and immunoglobulin genes^{1,2}, but the role of γ remains unknown. A central problem is to understand the conditions under which γ RNA is expressed in cells. The transcription of γ is abundant in T cells of fetal thymus³, but is negligible in peripheral T cells of adults, suggesting that γ is involved in development of the T-cell repertoire³. However, γ RNA was originally cloned from established lines of cytotoxic T cells (CTLs) derived from adult mice^{1,4} and this expression has been ascribed to non-physiological cell growth³. Possibly consistent with this, most of the γ RNA derives from genes rearranged abortively at the V_γ - J_γ junction of immunoglobulin genes, where V is the variable segment and J the joint segment (refs 5, 6; see ref. 7 for review). Here, we report the detailed analysis of γ transcription in T cells of adult mice, and find that transcription may occur in T cells with a broad range of surface phenotypes; that it is predominantly of a single V_γ - C_γ unit (where C is the constant region); and that in cells freshly explanted from animals it can be of productively rearranged genes.

We first examined whether γ -gene expression in cells of adult mice is restricted to established cell lines and hybridomas. Aliquots of $\sim 10^8$ T cells purified from spleens of 6-week-old BALB/c mice were either collected (NT cells) or were placed into short-term culture and subjected to one of three proliferative stimuli: lectin; allogeneic mouse cells; and anti-Thy-1 antibodies (see Fig. 2 legend). Extensive blast formation and proliferation occurred in each case. After 2 to 5 days, these cultures were collected for cellular nucleic acids. Each RNA (15 μ g) was electrophoresed, transferred to nitrocellulose and probed sequentially with four nick-translated DNAs specific for: (1) the unique C-region gene of the murine T-cell receptor (TCR) α -locus; (2) the two murine TCR β -chain C-regions; (3) any of the three homologous γ -chain C-regions, one of which, C7.5, is a pseudogene²; and (4) the *c-myc* proto-oncogene. All samples express equivalent amounts of TCR α - and β -chain RNAs but, as noted previously³, NT cells and d2 cells (T cells subjected to 2-day lectin stimulation) express only negligible γ RNA (Fig. 1). This applies equally to d3 and d5 cells (T cells maintained

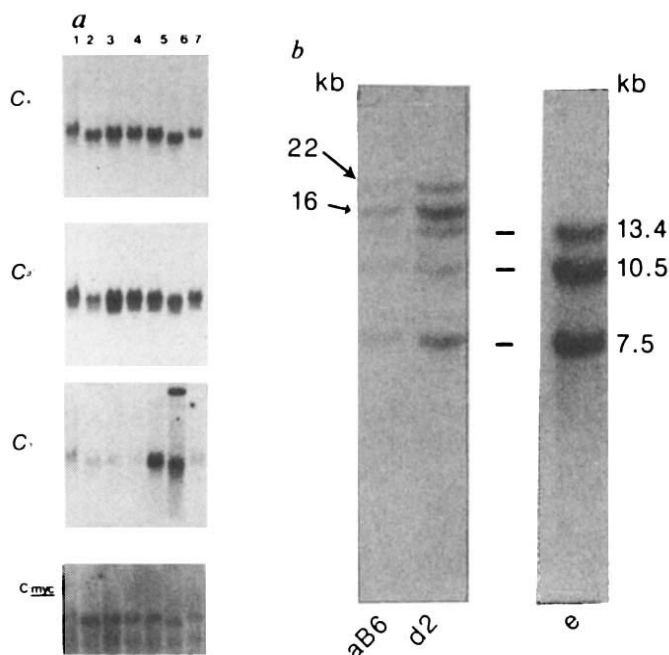


Fig. 1 a, Northern analysis of total cellular RNA derived from lane 1, NT cells; 2, d2 cells; 3, d3 cells; 4, d5 cells; 5, aB6 cells; 6, aDBA/2 cells; 7, T cells stimulated with a rabbit anti-mouse brain antibody that is directed against Thy-1 (ref. 9). Probes: C α , derived from pHDS58 (ref. 19) includes the last two exons and the 3' untranslated region of the unique murine C α gene¹⁷; C β , derived from pHDS11 (ref. 1), is an internal *NcoI* fragment from murine C β 2 that is also very similar to murine C β 1; C γ , derived from pHDS203 (ref. 1), is an *AvaI*-*PstI* fragment encompassing the *J*-C γ 10.5 (ref. 2) homologous to the equivalent regions of C γ 7.5 and C γ 13.4 (ref. 2); c-myc, a *Clal*-*EcoRI* fragment encompassing exon 3 of human c-myc²⁰. Use of synthetic size markers shows all major RNA species to be in agreement with previously reported sizes. The band below that of c-myc represents a remnant of the C β probing that directly preceded it and that was not stripped. b, Southern analysis of *EcoRI*-cleaved DNA from aB6 cells, d2 cells and BALB/c embryos (e) probed with the C γ fragment. Sizes in kilobases (kb) were determined from co-electrophoresis of ³²P-labelled λ DNA cut with *HindIII*, and are in agreement with the sizes published for germ-line and rearranged BALB/c γ genes². RNA was prepared by guanidinium thiocyanate-Sarkosyl lysis of cells, and repeated extractions with hot phenol and treatments with proteinase K. Electrophoresis and transfer were as published²¹. Exposure times were 2 days at -70 °C.

in Concanavalin A (con A) plus interleukin-2 for 3 or 5 days). By contrast, γ RNA is present in BALB/c cells stimulated for 5 days with cells of either C57BL/6 mice or DBA/2 mice. BALB/c and C57BL/6 mice differ at the major histocompatibility locus (MHC): BALB/c mice are *H-2^d*; C57BL/6, *H-2^b*. DBA/2 mice are also *H-2^d*, but differ from BALB/c mice at the Mls locus on chromosome 1. Cells of both strains elicit a proliferative mixed lymphocyte reaction (MLR) with T cells of BALB/c mice⁸ that is thought to result from direct engagement of alloreactive TCRs on the responding T cells. However, direct engagement of Thy-1, another cell-surface protein that triggers proliferation⁹ of the responders, does not produce a population of cells expressing the γ -chain gene (Fig. 1, lane 7). The radically different levels of γ RNA cannot easily be related to the status of the γ -chain genes: there seems to be an equivalent degree of rearrangement in d2 cells and the anti C57BL/6 (aB6) cells (Fig. 1b). Also, there is no correlation between γ -gene expression and that of c-myc, a proto-oncogene regarded as an index of T cells in active cycling¹⁰ (Fig. 1).

FACS (fluorescence-activated cell sorter) analysis of these cell populations (Fig. 2) shows that all are Thy-1⁺ and so the

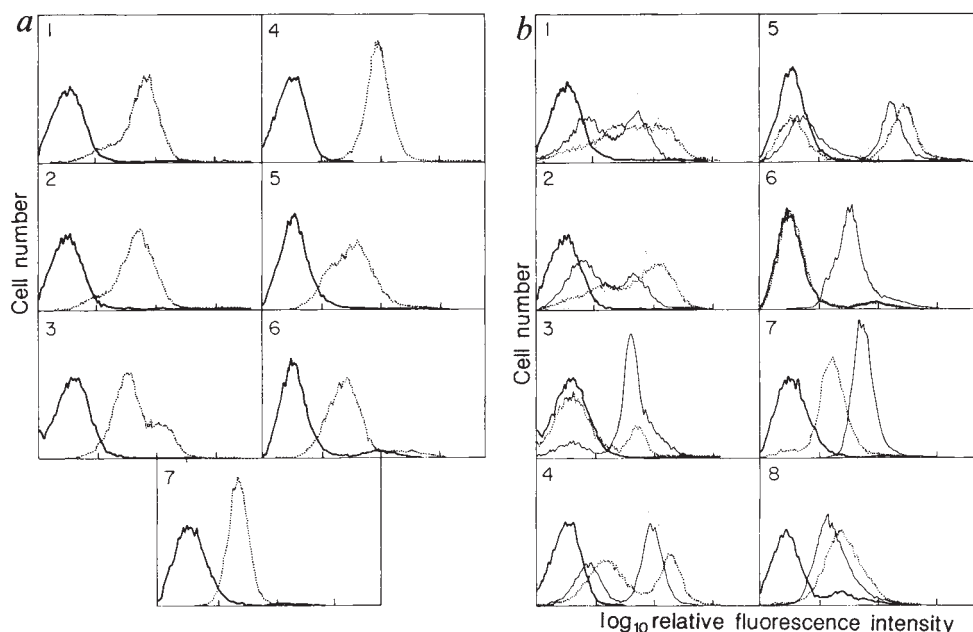


Fig. 2 FACS analysis of the surface phenotypes of the cell populations from which nucleic acid was prepared. *a*, Cells stained with the anti-Thy-1 monoclonal antibody Y-19 (ref. 9) (.....). *b*, Cells stained with monoclonal antibodies: anti-L3T4²⁵ (—); anti-Lyt 2 (Becton Dickinson) (.....); or a mixture of equal parts of each (.....). Controls (—) samples in which the biotinylated antibodies were omitted. Cell samples: 1, aB6; 2, aBr; 3, aDBA/2; 4, Con A d5; 5, RaMBR d5; 6, D10; 7, 5.5; and 8, 5.9. 1×10^6 cells were stained using biotin-conjugated antibodies and a 'second layer' of fluoresceinated avidin (Vector) as previously described²⁴.

Methods. BALB/c (H-2^d, Mls^b) T cells were purified from splenic cells by passage over an immunoglobulin anti-immunoglobulin-coated glass bead column²² and stimulated in 5-day

MLR cultures with allogeneic mitomycin C-treated spleen cells from C57BL/6 (H-2^b), B10.Br (H-2^k) or DBA/2 (H-2^d, Mls^a) strains. The cultures have been described in detail previously²³; they contained $2-4 \times 10^6 \text{ ml}^{-1}$ T cells responding to $4 \times 10^6 \text{ ml}^{-1}$ stimulator cells. In the mitogen-stimulated cultures, mitomycin C-treated BALB/c spleen cells replaced the allogeneic stimulators, and either $2.5 \mu\text{g ml}^{-1}$ Con A or 1/200 dilution rabbit anti-mouse brain (RaMBR) antiserum⁹ was added as a polyclonal mitogen in the presence of 5–10% supernatant from Con A-stimulated rat spleen cells added as a source of interleukin-2 (ref. 13). The AKR mouse T-cell clone D10 was propagated *in vitro* by weekly stimulation with the antigen conalbumin and B10.Br feeder cells in interleukin-2-containing medium¹³. The BALB/c T-cell clones 5.5 and 5.9 were grown under similar conditions in the presence of the antigen ovalbumin and BALB/c feeder cells¹².

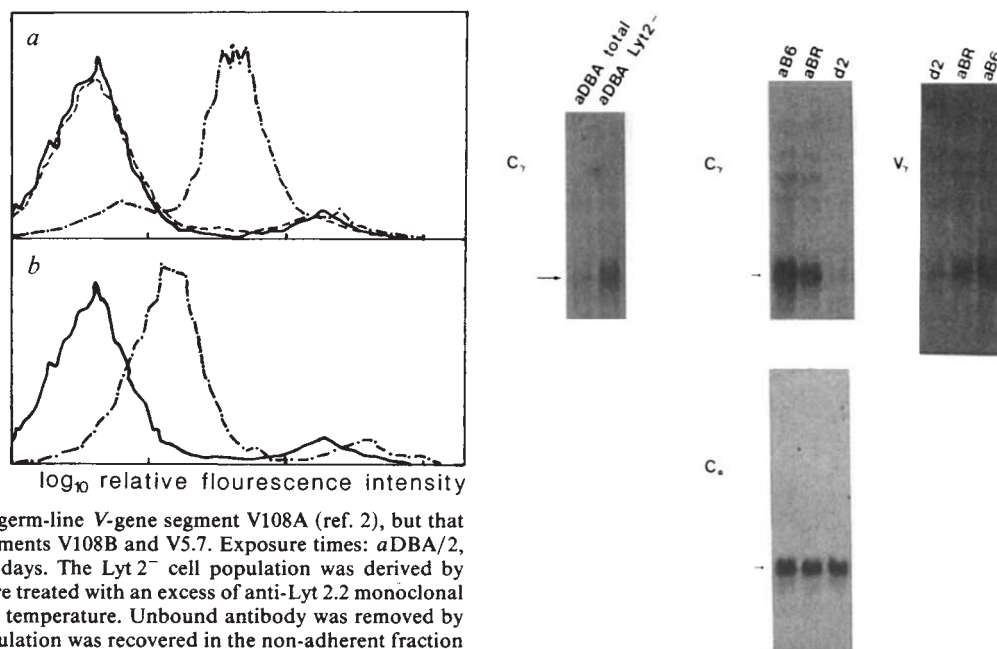
negligible γ expression in, for example, d5 cells cannot be attributed to poor purification of T cells or their subsequent dilution by lectin activated B cells. Cells 'double-negative' (DN) for Lyt 2 and L3T4 express γ in adult mice³, but such cells are a very minor component of any of the populations examined here (Fig. 2), implying that γ transcription in aB6 and aDBA/2 cultures derives from mature T cells that are Lyt 2⁺ and/or L3T4⁺. However, there is insufficient excess of Lyt 2⁺ or L3T4⁺ cells in the aB6 cells relative to the d5 cells to be certain in assigning γ RNA expression to either subset.

After stimulation for 5 days with Mls^a-bearing DBA/2 spleen cells, the blasting BALB/c T cells, which express γ , are mainly L3T4⁺, Lyt 2⁺ (Fig. 2). An uncloned line of these cells was

established (Fig. 3 legend) and from this a population of Lyt 2⁺ cells (Fig. 3a) was derived by negative selection. This population expresses γ RNA at higher levels than the unfractionated line (Fig. 3c), showing that T-cell γ RNA is not restricted to the Lyt 2⁺ subset.

Also in Fig. 3c are analyses of RNA from aB6, d2 and aBR cells (BALB/c T cells stimulated for 5 days with mitomycin C-treated spleen cells of B10.BR mice). The existence of RNA in aBR cells implies that it can be derived regardless of stimulating haplotype. Furthermore, in aB6 and aBR cells, the γ RNA is mainly, if not wholly, reactive to a V_γ probe excised from the germ-line V_γ gene segment that is rearranged and expressed in the BALB/b CTL, termed 2C. This V_γ segment (V109A; ref. 2)

Fig. 3 *a, b*, FACS analyses of an uncloned line of aDBA/2 cells, negatively selected for Lyt 2⁺ cells. 1, Cells stained with anti-Lyt 2 (—), or anti-L3T4 (.....). 2, Cells stained with anti-Thy-1 monoclonal antibody Y19 (.....). Controls (—) samples in which the biotinylated antibodies were omitted. The anti-Lyt 2 antibodies show a conspicuous fleck of activity at high log-fluorescence value, but the coincidence of that with the controls implies strongly that this fleck is attributable to dead cells. *c*, Northern analyses of total RNA performed as described in Fig. 1 legend. V_γ probe, a *Pst*I–*Ava*I fragment of pHDS203 (ref. 1) that derives from germ-line V_γ -gene segment V108A (ref. 2), but that is also homologous with V_γ -gene segments V108B and V5.7. Exposure times: aDBA/2, total; aDBA/2, 10 h; remainder, 2 days. The Lyt 2⁺ cell population was derived by negative selection. Cells 1.5×10^7 were treated with an excess of anti-Lyt 2.2 monoclonal antibody SR-147 for 30 min at room temperature. Unbound antibody was removed by washing three times. The Lyt 2⁺ population was recovered in the non-adherent fraction after panning of the antibody-coated T cells on anti-mouse-Ig-coated dishes²¹.



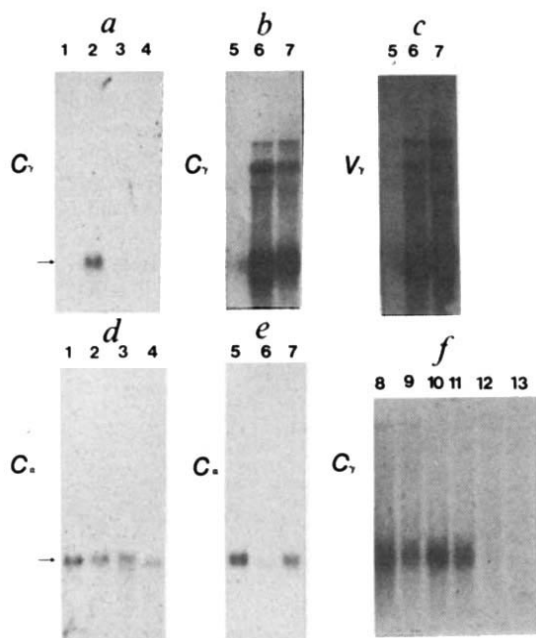


Fig. 4 Northern analysis of 15 µg total (t) or 100 ng poly(A⁺) (A) RNA from the following sources: lanes 1, 5, D10(t); 2, 6, 5.9(t); 3, Df1.12(t); 4, 5.5 (A); 7, 5.5(t); 8, 5.9(t); 9, 5.9(t); 10, 5.5 grown with ova (t); 11, 5.5 grown with Con-A(t); 12, D10(t); 13, NT cells (t). Exposure times: a, d, e, 12 h; b, c, 5 days; f, 9 h.

is similar to two other V_γ segments, V108B and V5.7, that comprise a V_γ family, V_γ2C. Therefore, stimulation with different haplotypes yields responder T cells that express V_γ gene segments from the same family.

Other reports^{4,11} contrast frequent transcription of γ in lines of mature Lyt-2⁺ CTLs restricted to class I MHC antigens, with only rare transcription in L3T4⁺ class II MHC-restricted T-helper cell lines. Therefore, the continued expression of γ RNA is the Lyt 2⁻ Mls population (Fig. 3c) prompted us to examine 5.5, 5.8 and 5.9, all T-helper cell lines derived from lymph nodes of BALB/c mice immunized with ovalbumin (ova)¹², and D10, derived from an AKR mouse immunized with conalbumin¹³. All cell lines display 'classical' T-helper phenotypes, for example, they are L3T4⁺ and they help B cells respond to specific antigen in the context of class II MHC products (I-a^d for 5.5, 5.8 and 5.9; I-a^k for D.10). However, there is dramatic disparity in C_γ expression (Fig. 4a,b). Like other T-helper cells, D.10 expresses little C_γ (ref. 11), but in contrast, 5.5 and 5.9 cells express more γ RNA than the positive control, class I-reactive CTL Df1.12. For these cells, the expressed V_γ is of the V_γ2C family (Fig. 4c). The T-helper cells (5.5, 5.8 and 5.9) that express γ RNA are also positive for Lyt 2 (Fig. 2), whereas D.10 cells are not. However, the expression of γ in the aDBA/2 line indicates that Lyt 2 expression is not an obligatory correlate of γ-chain expression.

Like many cloned T-cell lines, 5.5 cells are maintained in the presence of antigen. Because we have found γ RNA in BALB/c cells in response to allogeneic stimuli thought to act via the TCR, we considered that the possibility that the high levels of γ RNA in 5.5 cells is maintained by antigen-TCR interaction. We found that 5.5 cells grown for >3 weeks in an interleukin-2-rich medium containing Con-A expressed unaltered levels of γ RNA instead of antigen (Fig. 4f). The same result was obtained using 5.9 cells (data not shown).

To elucidate further γ-chain gene expression in cells of adult mice freshly explanted to culture, we constructed a complementary DNA library to aB6 cell RNA (Fig. 5 legend) and screened it with a C_γ probe. From ~2 × 10⁵ independent clones six were

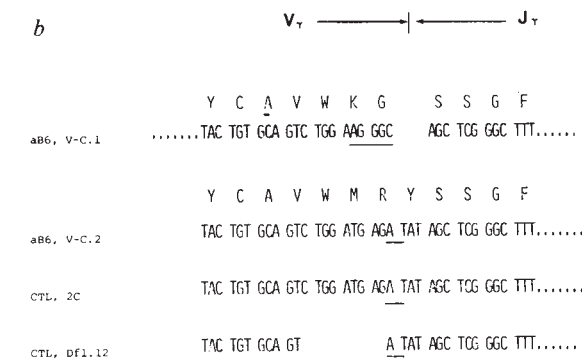


Fig. 5 a, DNA sequence of a prototypic (one of six) cDNA from aB6 cells compared at its 3' end with the sequences of BALB/c germ-line C segments, 10.5 and 7.5 (ref. 2), and 13.4 (determined *de novo*). Asterisks, mismatches to the cDNA. b, The V-J junctions of two aB6 cDNA clones (clones V-C1, and V-C2) and the V-J junctions of γ cDNAs derived from CTLs 2C and Df1.12 (refs 1, 4). Nucleotides of unknown origin are underlined. The cDNA clones were constructed from poly(A)⁺ RNA and ligated to λgt10 arms according to Huynh *et al.*²⁶. Fragments labelled with ³²P using T4 kinase or the Klenow fragment of *Escherichia coli* DNA polymerase I were sequenced according to ref. 27.

detected, sequenced in the C region and found to be identical. A sequence of a prototypic cDNA (Fig. 5a) is compared with sequences from the 3' end of C_γ 10.5 (ref. 2), C_γ 7.5 (ref. 2) and C_γ 13.4 (determined *de novo*). In this 3' region, the three C gene segments are readily distinguishable (Fig. 5a); the clones derive from the C_γ 10.5 gene segment, which is the segment expressed in 2C CTLs¹.

Two of the six cDNA clones extended into the V gene segments, where again, comparison to published germ-line sequences² allowed assignment of both cDNAs to V108A, the V gene segment used in 2C cells. Furthermore, analysis of the V-J junction in these clones shows two out of two to be productive rearrangements (Fig. 5b). In-frame joining in γ-genes of CTLs 2C and Df1.12 occurs by insertion of an AT dinucleotide between V and J^{2,4}. (There are also nucleotides deleted in the Df1.12 join.) The productive rearrangement of aB6 cDNA, V-C2, also occurs by an AT insertion: in fact the junction is identical to that in 2C cells¹. By contrast, aB6 clone V-C1 shows deletion of six nucleotides and insertion of five. These five may be template-independent insertions. However, their sequence, A(G)₃C is strikingly similar to A(G)₃C, which is the sequence at the 3' end of the first D (diversity) gene segment in murine

TCR β -chain genes^{14,15}. A similar homology was used previously to predict the existence of *D* elements in the murine TCR α -chain locus¹⁶ and, also in that locus, to propose a germ-line *D-J* fusion event¹⁷.

We have shown here several examples of γ -RNA expression in adult mouse T cells. Taking 5.5, 5.8 and 5.9 cells (L3T4⁺, Lyt 2⁺) and aDBA/2 cells (L3T4⁺, Lyt 2⁻), together with CTLs 2C and Dfl.12 (Lyt 2⁺) and adult peripheral DN cells (refs 1, 8, 9), we conclude that adult cells with a broad range of cell-surface markers transcribe γ -chain genes. Significant levels of γ RNA cannot be induced in short-term culture by Con A plus interleukin-2, although cells stimulated in this manner contain rearranged γ -chain (Fig. 1b). Hence, rearrangement is not sufficient for γ -gene expression. By contrast, γ RNA is elicited

in cells subject to various allogeneic stimuli. We do not yet know the properties of the cells expressing this RNA, nor whether they are activated directly or indirectly by the stimuli. But it is striking that our protocol has yielded transcripts of functional *V-J* rearrangements from adult immunocompetent mice that until now were the exception rather than the rule (refs 5, 6, 18; see ref. 7 for review). Finally, one of these productive rearrangements is the same as that previously derived from 2C, an established CTL line.

We thank B. Broughton, P. Conrad, E. Mjolsness and Shirley Smith for help. Supported by NIH grants AI-14579, and CA-29606 to C.J. Jr and NIH grant CA-40364-01 to A.H. C.J. Jr and B.J. were supported by the Howard Hughes Medical Institute; A.H. is a Searle Scholar.

Received 4 April; accepted 30 July 1986

1. Saito, H. *et al.* *Nature* **309**, 757-762 (1984).
2. Hayday, A. C. *et al.* *Cell* **40**, 259-269 (1985).
3. Raulat, D. *et al.* *Nature* **314**, 103-107 (1985).
4. Kranz, D. M. *et al.* *Nature* **313**, 752-775 (1985).
5. Rupp, F. *et al.* *Nature* **321**, 876-878 (1986).
6. Reilly, E. B., Kranz, D. M., Tonegawa, S. & Eisen, H. N. *Nature* **321**, 878-880 (1986).
7. Gelfer, M. & Marrack, P. *Nature News and Views* **321**, 116 (1986).
8. Katz, M. E. & Janeway, C. A. *J. Immun.* **134**, 2064-2070 (1985).
9. Jones, B., *Eur. J. Immun.* **13**, 678-684 (1983).
10. Kelly, K. *et al.* *Cell* **35**, 603-610 (1983).
11. Heilig, J. S. *et al.* *Nature* **317**, 68-70 (1985).
12. Conrad, P. M. & Janeway, C. A. *Immunogenetics* **20**, 311-319 (1984).
13. Kaye, J. *et al.* *J. exp. Med.* **158**, 836-856 (1984).
14. Kavalier, J., Davis, M. M. & Chien, Y.-H. *Nature* **310**, 421-423 (1984).

15. Siu, G. *et al.* *Nature* **311**, 344-350 (1984).
16. Chien, Y.-H. *et al.* *Nature* **312**, 31-35 (1984).
17. Hayday, A. C. *et al.* *Nature* **316**, 828-832 (1985).
18. Quertermous, T. *et al.* *Nature* **322**, 184-187 (1986).
19. Saito, H. *et al.* *Nature* **312**, 36-39 (1984).
20. Hayday, A. C. *et al.* *Nature* **307**, 334-340 (1984).
21. Bond, J. F. & Farmer, S. R. *Molec. cell. Biol.* **3**, 1333-1342 (1984).
22. Jones, B. & Janeway, C. A. *Eur. J. Immun.* **11**, 584-592 (1981).
23. Janeway, C. A., Lerner, E. A., Jason, J. M. & Jones, B. *Immunogenetics* **10**, 481-497 (1980).
24. Jones, B., Tite, J. P. & Janeway, C. A. *J. Immun.* **136**, 348-356 (1986).
25. Dialynis, D. P. *et al.* *Immun. Rev.* **74**, 29-56 (1983).
26. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A practical Approach* (ed. Glover, D. M.) 49-78 (IRL, Oxford, 1984).
27. Maxam, A. & Gilbert, W. *Methods Enzym.* **65**, 499-560 (1980).

A unique T-cell receptor complex expressed on human fetal lymphocytes displaying natural-killer-like activity

P. Moingeon*, A. Ythier*, G. Goubin†, F. Faure*, A. Nowill*, L. Delmon*, M. Rainaud‡, F. Forestier‡, F. Daffos‡, C. Bohuon* & T. Hercend*§

* Unité de Biologie Cellulaire, Institut Gustave-Roussy, rue Camille Desmoulins, 94805 Villejuif Cedex, France
† Laboratoire d'Oncogénèse Moléculaire, Institut Curie, rue d'Ulm 75005, Paris, France

‡ Centre de Diagnostic Prénatal de de Foetologie, Hôpital Notre-Dame de Bon Secours, rue Giordano Bruno, 75014 Paris, France

We have recently derived a series of cloned cell lines¹ displaying natural killer (NK) cell-like activity from normal human fetal blood (25 weeks)². The lines were obtained after repeated stimulation of mononuclear cells with allogeneic Epstein-Barr virus (EBV)-transformed B lymphocytes and are interleukin-2 (IL-2) dependent. Initial characterization of the clones has been reported previously¹. Certain of these clones have been found to have unusual surface characteristics, namely, they are recognized by several well-defined anti-T3 antibodies, but do not react with WT31 (ref. 3), which is thought to recognise an invariant epitope of the human (Ti- $\alpha\beta$) structure^{4,5}. Transcription of the genes encoding the α - and β -chains of the T-cell receptor was assessed in two of these clones (F6A4 and F6C7). Ti- β genes were found to be expressed, whereas α messenger RNA was not detected in Northern blot analysis. These data strongly suggest that these cells do not produce a stoichiometric T3/Ti- $\alpha\beta$ receptor complex. However, experiments performed with a monoclonal antibody (anti-NKFi) developed against F6C7 cells demonstrated the existence of a unique clonotypic structure [relative molecular mass (*M*_r) 85,000 (85K)] which is surface-associated with T3 proteins. Furthermore, both anti-T3 and anti-NKFi were found to block cytotoxic effector

function. Together, the results support the view that T3 proteins are involved in non-major histocompatibility complex (MHC)-restricted cytotoxic reactions mediated by certain circulating fetal lymphocytes which are likely to use a clonotypic structure distinct from both the 'first' ($\alpha\beta$)⁴ and the putative 'second' ($\gamma\delta$)⁶ T-cell receptor to recognize their target. The present studies were designed to characterize this structure.

Clone F6C7, which exhibits a strong proliferative response to IL-2, was selected for these experiments. As previously described, F6C7 cells are T3⁺, WT31⁻, T8⁺, T11⁺ and NHK1A⁺; they kill conventional NK target cell lines such as K562, MOLT-4 and JM in an MHC-independent fashion.

We first studied the expression of Ti genes in the fetal clone F6C7, using a complementary DNA which hybridizes both to constant region β 1 and β 2 segments. Two distinct Ti- β RNA species of 1.0 and 1.3 kilobases (kb) were detected in F6C7 cells (Fig. 1, lane g). The ratio between full-length (1.3 kb) and truncated (1.0 kb) mRNA was lower in the fetal lymphocytes than in control JM leukaemia cells (Fig. 1, lane f). For Ti- α RNA, a variable-joining-constant region cDNA clone was used. A signal was observed at 1.6 kb in both JM (Fig. 1, lane a) cells and in two conventional IL-2-dependent cloned T-cell lines used here as additional controls (lanes b, c). There was no corresponding band when probing RNA from F6C7 cloned lymphocytes (Fig. 1, lane d). To exclude misinterpretation, the same blot was dehybridized and rehybridized with the β probe. This procedure allowed the identification of the 1.0- and 1.3-kb bands described above (not shown).

Taken together with the absence of WT31 monoclonal antibody reactivity, these findings strongly suggest that F6C7 cells do not express a conventional stoichiometric (T3/Ti- $\alpha\beta$) receptor complex. Anti-T3 immunoprecipitations following chemical cross-linking were then used to assess whether a molecule distinct from Ti- $\alpha\beta$ was associated with T3 proteins on the surface of these fetal lymphocytes. As expected, two bands at 20K and 27K were found in SDS-polyacrylamide gel electrophoresis (PAGE) analysis after precipitation with anti-T3 (Fig. 2, lane a). More importantly, incubation of the cells with the crosslinking agent dithiobis propionic acid *N*-hydroxysuccinamide ester (DSP) before immunoprecipitation led to the detection (after reduction with 2 mercaptoethanol) of an additional 44K band (Fig. 2, lane b).

§ To whom correspondence should be addressed.

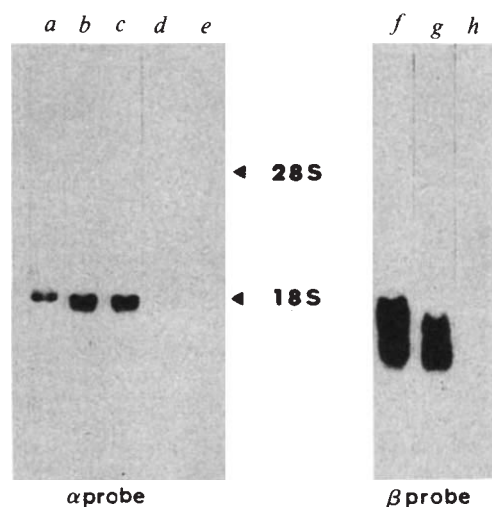


Fig. 1 Northern blot analysis of F6C7 cloned cell lines with α (lanes a–e) and β (lanes f–h) Ti probes. Lanes: a, f, JM cells; b, c, two control IL-2-dependent T-cell clones; d, g, F6C7; e, h, EBV-transformed B-lymphoblastoid cell line LAZ 461.

Methods. RNA was prepared from the cytoplasmic fraction of cells lysed with Nonidet P-40 and centrifuged through a sucrose cushion to remove nuclei, followed by proteinase K digestion, phenol and chloroform extraction, and ethanol precipitation (the last procedures being repeated twice). RNA preparations were dissolved in sterile water, heated in the presence of formamide and formaldehyde at 65 °C for 5 min, and run in a 1% agarose gel containing 2.2 M formaldehyde. After transfer to nitrocellulose membranes, hybridizations were carried out in 5 ml of 50% formamide, $5 \times \text{SSC}$, $1 \times \text{Denhardt's}$, 10% dextran sulphate with $1-2 \times 10^6$ c.p.m. ml^{-1} of ^{32}P -labelled α (lanes a–e) or β (lanes f–h) Ti probes for 2 h at 42 °C. The membranes were washed twice in $1 \times \text{Denhardt's}$, $3 \times \text{SSC}$ at room temperature for 30 min, then three times in $0.1 \times \text{SSC}$, 0.1% SDS at 50 °C for 45 min. SDS was removed by several washes of $0.1 \times \text{SSC}$ at room temperature. The Ti- α probe was derived by digesting plasmid pGA5 (provided by Dr T. Palmer) with *EcoRI* to yield a fragment of ~1.2 kb representing a complete Ti α -chain transcript. The Ti- β probe was derived by digesting plasmid pBDOB1 (provided by Dr T. Mak) with *BglIII* to yield a fragment of ~450 bases containing sequences of the constant region of Ti β -chain. The inserted DNA fragments were separated on 1% agarose gel, run onto DEB1 paper (Whatman), eluted, labelled by nick translation and 15- μg aliquots of RNA were run in each lane.

F6C7 cloned cells were used to immunize mice in order to generate monoclonal antibodies directed against this putative 44K molecule. Initial screening of hybridomas was based on the ability of their supernatants to block cytotoxic activity of F6C7 lymphocytes against K562 cells. Out of several positive hybridomas one cell line, termed NKFi, was studied further because its supernatant did not block the cytotoxic activity of JT9 cloned NK lymphocytes which express NKTa, a previously described Ti- $\alpha\beta$ clonotypic determinant⁸. The NKFi hybridoma was recloned twice by limiting dilution and injected into mice to generate immune ascites. The anti-NKFi monoclonal antibody (IgG₁ subclass) was tested by indirect immunofluorescence against a series of selected target cells. NKFi was not expressed on either peripheral blood lymphocytes or day 7 phytohaemagglutinin-stimulated blasts from three normal adult donors (not shown). In addition, anti-NKFi was totally unreactive with a variety of haematopoietic cell lines either of T-cell origin such as JM, MOLT-4 and CEM, or non-T-cell origin, including LAZ388, LAZ156, LAZ461, LAZ509 (EBV-transformed lymphoblastoid lines), K562, U937, KGI and HL60 (tumour cells of myelo-monocytic derivation). Expression of NKFi was tested both on the cell line from which F6C7 was derived and on several clones obtained simultaneously. In the bulk culture (which was frozen when cloning was performed) less than 5%

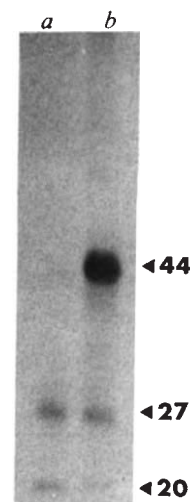


Fig. 2 Immunoprecipitations with anti-T3 monoclonal antibody. Immunoprecipitations were performed either directly (lane a) or following chemical crosslinking (lane b). F6C7 cells were surface labelled with ^{125}I using a standard lactoperoxidase method described previously⁴. Following several steps of preclearing with *Staphylococcus A* suspension as well as irrelevant antibodies, specific precipitation was carried out for 6 h at 4 °C with anti-T3 antibody coupled to protein A-Sepharose-4B beads. SDS-PAGE was performed under reducing conditions (with 5% 2-mercaptoethanol). DSP (Sigma) was used for chemical crosslinking as described by Brenner *et al.*⁷. Briefly, cells were incubated at 5×10^6 cells ml^{-1} in phosphate-buffered saline containing a final concentration of $50 \mu\text{g ml}^{-1}$ DSP for 1 h at room temperature. Cell lysate was then prepared and DSP was removed by dialysis before carrying out precipitations using the procedure described above.

of the cells reacted with anti-NKFi. Some of the clones expressed NKFi, suggesting that the same cell was cloned several times. Perhaps more importantly, several previously described¹ T3⁺ WT31⁺ T8⁺ NKH1A⁺ cloned cell lines displaying NK-like activity, such as F6A4, F6A7, F6D7, F6F8 and F6G7, were found to be NKFi⁺ (note that RNA extracted from F6A4 was also tested in Northern blot analysis (not shown); results were similar to those obtained for F6C7: both 1.0- and 1.3-kb β mRNA was present without detectable transcription of α). Taken together, these data demonstrate that NKFi is a surface structure with clonotypic cell distribution.

Co-modulation experiments were conducted to demonstrate membrane association between NKFi and T3 proteins. A 24-h incubation of F6C7 cells with a saturating concentration of anti-NKFi resulted in a strong decrease of both NKFi and T3 expression whereas it had no effect on the reactivity of anti-T11 used here as control (Fig. 3). Reciprocal experiments in which anti-T3 was used to induce modulation gave similar results (Fig. 3). Biochemical characterization of the structure recognized by anti-NKFi was performed by SDS-PAGE analysis following specific immunoprecipitation from ^{125}I -radiolabelled F6C7 cells. NKFi was found to resolve under non-reducing conditions (Fig. 4, lane b) as a single band of M_r 85K; after addition of 2-mercaptoethanol, a major band was detected at 44K, while a faint band was always present at 41K (Fig. 4, lane d). Control immunoprecipitations under identical conditions led to the expected identification of the NKTa Ti- $\alpha\beta$ clonotypic determinant⁸ expressed by JT9 cells (95K unreduced; Fig. 4, lane a; 49–44K reduced, lane c).

Previous studies^{8,9} strongly suggest a direct involvement of the NKTa Ti- $\alpha\beta$ receptor in the NK-like activity of certain T-cell clones derived from adult peripheral blood. It was important here to demonstrate that the T3/NKFi complex was functional and involved in cytotoxic reactions mediated by the fetal cloned cells. Blocking experiments were performed to address this point. Table 1 shows that treatment of F6C7 lymphocytes

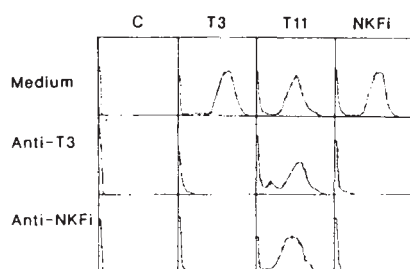


Fig. 3 Co-modulation of T3 and NKFi antigens. F6C7 cells were divided into three groups (vertical axis) and incubated for 24 h with medium, anti-T3 or anti-NKFi at saturating concentrations. Cells from each group were then washed three times and retested at 4 °C for expression of T3, T11 and NKFi antigens (horizontal axis) using an indirect immunofluorescence assay. Each population was analysed using an Epics C flow cytometer. Anti-NKFi was used as a negative control (C) in these experiments.

by either anti-T3 antibody or anti-NKFi antibody resulted in >55% and 74% inhibition of cytotoxicity against K562 and JM cells respectively, whereas anti-NKFi and anti-T8 (used as control reagents) had no inhibitory effects.

Taken together these data support the view that certain T lymphocytes circulating in a 25-week-old normal human fetus express a unique surface receptor which has not been previously identified. NKFi is not an α/β heterodimer⁴. Furthermore, its biochemical structure is distinct from that of the non-disulphide-linked $\gamma\delta$ complex which has been very recently proposed as a putative second T-cell receptor⁶. A likely possibility is that NKFi represents a $\beta\beta$ homodimer. However, preliminary experiments performed with anti- β F1 (ref. 6), a monoclonal antibody (kindly provided by Dr Michael Brenner) directed at a framework component of Ti- β did not provide evidence to substantiate this hypothesis. Indeed, we could not detect the β F1 antigenic determinant on the surface of WT31-F6C7 cells either by appropriate immunofluorescence assays or by immunoprecipitation. However, such negative results do not definitely exclude expression of a β -chain on these fetal lymphocytes because the β F1 epitope may be masked if the protein is associated with a polypeptide distinct from α . Thus, further studies are required to establish whether NKFi is a novel molecule or if it is made

Table 1 Effect of monoclonal antibodies on F6C7 cell cytotoxicity

Target cells	Monoclonal antibodies				
	Medium	Anti-NKFi	Anti-T3	Anti-T8	Anti-T8
K562	76	78	37	34	77
JM	68	64	15	18	63

Target cells were labelled with ⁵¹Cr and plated at 5,000 cells per well. Effector-to-target cell ratio was 10:1. Chromium release was evaluated following a 3-h assay and specific cytotoxicity calculated as described previously⁸. Effector cells were pretreated with monoclonal antibodies for 3 h. Saturating concentrations predetermined by indirect immunofluorescence were used. Numbers represent the mean specific cytotoxicity in triplicate cultures; s.d. $\leq$ 5%.

up of a novel combination of previously described T-cell receptor chains such as β , γ or δ . This structural characterization should contribute to an understanding of whether F6C7 cells are T lymphocytes arrested at an intermediate step of a major thymic differentiation pathway (for example T3⁺ WT31⁺ T4⁺ T8⁺ β mRNA⁺ F6C7 cells appear more 'mature' than the T3⁺ T4⁺ T8⁺ $\gamma\delta$ β mRNA⁺ thymocytes recently described by Bank *et al.*¹⁰) or if NKFi is expressed on cells from a related but distinct lineage. Because circulating F6C7 lymphocytes carry the NK-associated antigen NKH1A (ref. 11) and because they mediate unconventional non-MHC-restricted cytotoxicity, we favour the latter hypothesis. In any case these data, together

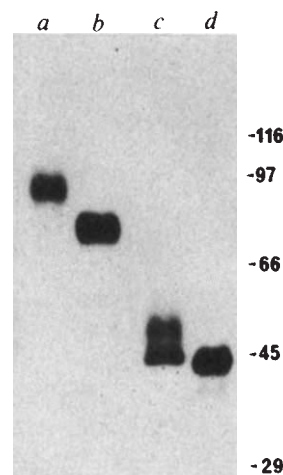


Fig. 4 Immunoprecipitation of NKFi structure. Immunoprecipitation experiments were performed (as Fig. 2 legend). Lanes: a, NKFi precipitated from JT9 cells, nonreducing conditions; b, NKFi precipitated from F6C7 cells, nonreducing conditions; c, NKFi, reducing conditions; d, NKFi, reducing conditions.

with the recent discovery of novel T3-associated polypeptides⁶, reveal additional levels of complexity in T-cell ontogeny. Hopefully, further defining the different receptor chains and their various possible associations will shed light on the lineage and the biological significance of those minor populations of T lymphocytes which display 'NK-like' activity.

We thank Dr Armand Bessus for discussions, Mrs M. Y. Guillard and E. Leclercq for technical assistance, and Mrs D. Rivierre and Miss V. Martel for preparation of the manuscript. This work was supported in part by ADRC grant N 6338 and CNRS contract N 06931. T.H. is a Special Fellow of the Leukemia Society of America.

Received 3 July; accepted 18 August 1986.

1. Nowill, A. *et al.* *J. exp. Med.* **163**, 1601-1606 (1986).
2. Desmonts, G. *et al.* *Lancet* **i**, 500-504 (1985).
3. Tax, W. *et al.* *Nature* **304**, 445-447 (1983).
4. Meuer, S. C. *et al.* *Nature* **303**, 808-812 (1983).
5. Spits, H. *et al.* *J. Immun.* **135**, 1922-1926 (1985).
6. Brenner, M. B. *et al.* *Nature* **322**, 145-149 (1986).
7. Brenner, M. B., Trowbridge, I. S. & Strominger, J. L. *Cell* **40**, 183-190 (1985).
8. Hercend, T. *et al.* *J. exp. Med.* **158**, 1547-1553 (1983).
9. Ritz, J. *et al.* *Science* **228**, 1540-1544 (1985).
10. Bank, I. *et al.* *Nature* **322**, 179-181 (1986).
11. Hercend, T. *et al.* *J. clin. Invest.* **75**, 932-943 (1985).

A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes

Louis M. Staudt*†, Harinder Singh*‡, Ranjan Sen*†, Thomas Wirth*†, Phillip A. Sharp*‡ & David Baltimore*†

* Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA
† Department of Biology and ‡ Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Immunoglobulin gene promoters are active only in lymphoid cells and this tissue-specific activity requires an octamer sequence, ATTTGCAT (refs 1-9). Paradoxically, this same octamer motif seems to be a transcriptional control element in promoters which are active in all tissues¹⁰. Using an electrophoretic mobility shift

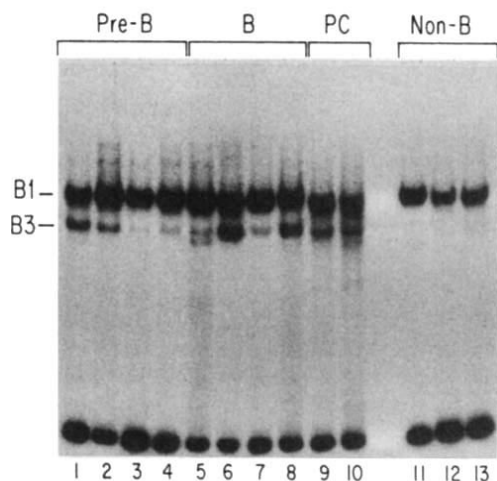


Fig. 1 Electrophoretic mobility shift assay of nuclear extracts from lymphoid and non-lymphoid cell lines. The nuclear extracts were prepared from the following cell lines: Lane 1, HAFTL: Harvey sarcoma virus transformant²³, which represents a very early stage in B-cell ontogeny because its germline H-chain locus is actively undergoing D_H-J_H rearrangements (S. Desiderio and D.B., unpublished data); lane 2, 18-81A-2: Abelson murine leukaemia virus transformant containing rearranged H-chain loci and germline κ -chain loci¹⁵; lane 3, PD: Abelson murine leukaemia virus transformant containing rearranged H-chain loci, which is actively rearranging its κ -chain loci¹⁹; lane 4, 70Z/3: mouse pre-B-cell line which contains rearranged H-chain and κ -chain loci but only transcribes its H-chain genes^{12,14,17,18}; lane 5, AJ/9: mouse Ly1-positive B-cell line²⁴; lane 6, WEHI-231: mouse B-cell line; lane 7, EW: human Epstein-Barr virus (EBV)-negative Burkitt's lymphoma cell line; lane 8, Raji: human EBV-positive Burkitt's lymphoma cell line; lane 9, SP2/0: mouse-mouse hybridoma; lane 10, S107: mouse plasmacytoma; lane 11, HeLa: human cervical carcinoma cell line; lane 12, L: mouse fibroblast cell line; lane 13, MEL: mouse erythroleukaemia cell line.

Methods. Nuclear extracts were prepared according to the method of Dignam *et al.*²⁵. The electrophoretic mobility shift assay was performed as described elsewhere¹¹ except that the gel running buffer was changed to 50 mM Tris, 380 mM glycine, 2 mM EDTA pH 8.5. The radiolabelled probe was an *EcoRI*-*PvuII* fragment of the plasmid pSPIgV κ containing the κ -promoter octamer sequence¹¹. Between 6 and 10 μ g of nuclear protein was used in the DNA binding reactions.

assay to identify DNA binding proteins, we have now detected two species of nuclear proteins which bind specifically to this octamer. One previously characterized form¹¹ (NF-A1) was found in all cell lines tested while the other form (NF-A2) was restricted to lymphoid cell lines. NF-A2 was found in cell lines representing all stages of B-cell differentiation and in half of the T-lymphoma cell lines tested. The identification of a lymphoid-specific octamer binding protein may account for the lymphoid-specific activity of immunoglobulin promoters.

Nuclear extracts were prepared from a variety of lymphoid and non-lymphoid cell lines and tested in the electrophoretic mobility shift assay using a radiolabelled κ -promoter fragment containing the octamer sequence motif (Fig. 1). All nuclear extracts displayed a band, B1, which is known to be generated by a ubiquitous octamer-specific DNA binding protein¹¹. The previously described weaker band, B2, which migrated slightly more slowly than B1, was also observed. Most importantly, a third band, B3, migrating faster than B1, was detected only in nuclear extracts from lymphoid cells (Fig. 1). This band was seen with nuclear extracts from cell lines representing all stages of the B-cell lineage, that is, pre-B, mature B and plasmacytoma. The B3-positive cell line HAFTL represents the earliest-defined stage in B-cell differentiation; it contains a germline immunoglobulin heavy (H)-chain locus which is spontaneously undergoing rearrangements in the heavy-chain diversity-joining

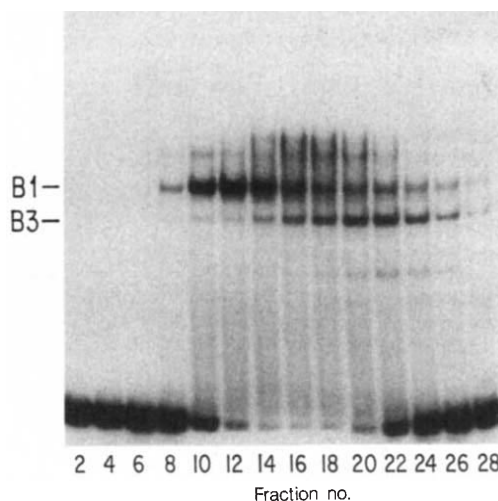


Fig. 2 Heparin-sepharose chromatography of WEHI-231 nuclear extract. WEHI-231 nuclear extract was prepared as described elsewhere²⁵ except that following the high salt elution of proteins from isolated nuclei with buffer C (20 mM HEPES pH 7.9, 0.42 M NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulphonyl fluoride, 25% v/v glycerol), the extract was passed over a DE52 column equilibrated with buffer C so as to remove nucleic acids. The effluent was dialysed against buffer D (20 mM HEPES pH 7.9, 0.04 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulphonyl fluoride, 20% v/v glycerol) and applied to a heparin-Sepharose 6CLB column equilibrated with the same buffer. Proteins eluted from this column with a linear 0.1–0.45 M KCl gradient in buffer D and alternate fractions analysed in the electrophoretic mobility shift assay (see Fig. 1).

region (D_H-J_H) (S. Desiderio and D.B., unpublished results). In addition, B3 was seen with nuclear extracts from two out of four of the T-cell lymphoma cell lines tested (data not shown). The level of B3 varied considerably between cell lines but independently prepared nuclear extracts from a given cell line always gave the same ratio of B1 to B3 (data not shown). The B-lymphoid cell line EW was particularly low in the factor giving rise to B3, explaining why we had previously detected only small amounts of it¹¹. Non-lymphoid cell lines which were found to be negative for B3 included HeLa, L, MEL, NIH 3T3, PCC4 and F9 (Fig. 1 and data not shown). Various minor bands were observed with non-lymphoid cell extracts, none of which co-migrated with B3 on extended electrophoresis (data not shown).

To determine whether B1 and B3 were generated by distinct protein components of the nuclear extract, the WEHI-231 nuclear extract was chromatographed on heparin-Sepharose and the fractions were tested in the DNA binding assay (Fig. 2). Using a linear KCl gradient elution, B1-generating activity was found to elute maximally in fraction 12 (0.2 M KCl) whereas B3-generating activity peaked in fraction 22 (0.35 M KCl). This ruled out the possibility that B1 and B3 were generated by the same protein binding to different conformations of the DNA probe. Thus, B1 and B3 seem to be generated by distinct octamer binding proteins, one lymphoid-specific and one ubiquitous, which can be separated biochemically. It is, however, quite possible that the two factors are related by post-translational modification such as proteolysis or phosphorylation, or are stable complexes of distinct factors with a common protein. We have previously referred to the ubiquitous factor as IgNF-A (ref. 11); we now propose to call the ubiquitous factor NF-A1 and the lymphoid-specific factor NF-A2.

To define precisely the contact residues on the DNA for the two factors, a methylation interference assay was used. For this assay, radiolabelled κ - or H-chain promoter fragments in which one of the two DNA strands was end-labelled were partially methylated by dimethyl sulphate and incubated with WEHI-231

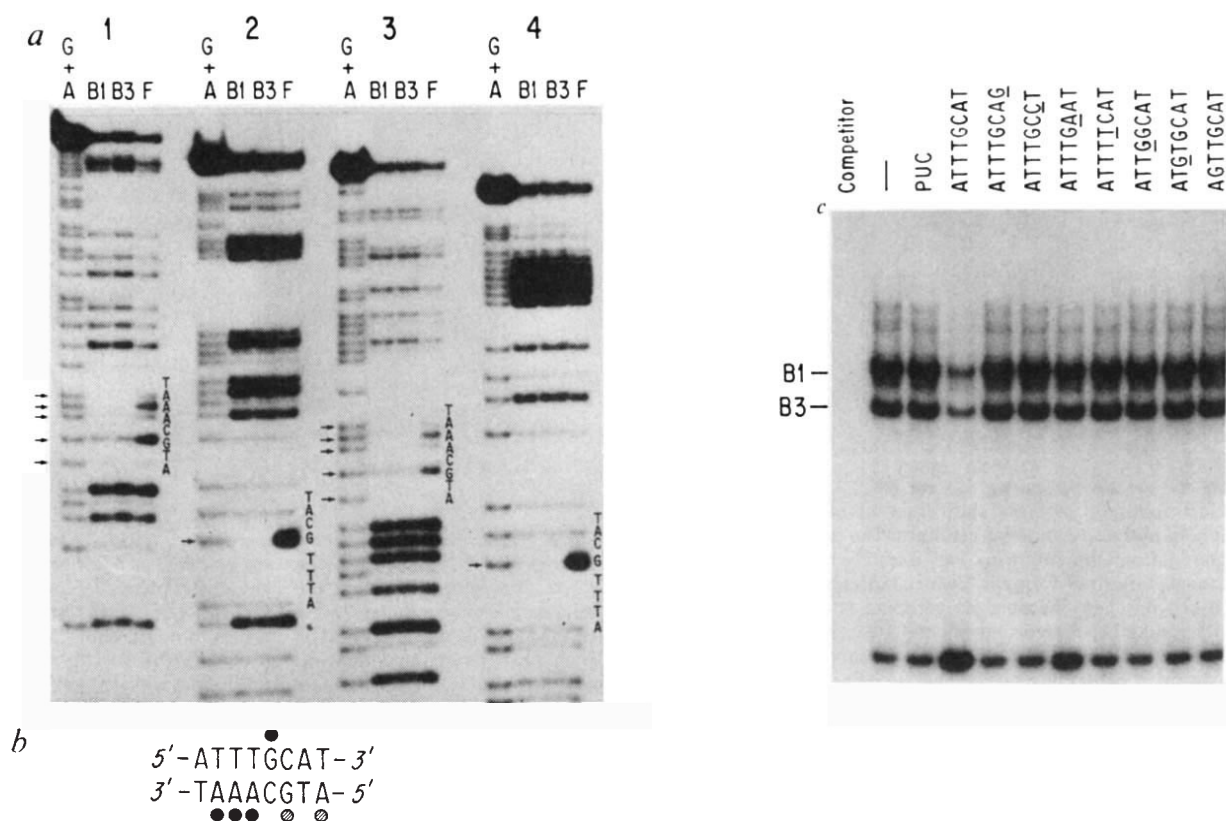


Fig. 3 *a*, Methylation interference analysis of octamer binding proteins. Panel 1, H-chain promoter coding strand; panel 2, H-chain promoter noncoding strand; panel 3, κ -chain promoter noncoding strand; panel 4, κ -chain promoter coding strand. Each panel contains a G+A reaction²⁶ of the probe DNA alongside the cleavage reactions from DNA in complexes B1, B3 and unbound DNA (F). *b*, Scheme of the methylation interference data. Full interference with the binding of the octamer binding proteins is shown as a closed circle and partial interference by a hatched circle. *c*, Fine-specificity analysis of octamer binding proteins.

Methods. DNA fragments of the H-chain promoter from the 17.2.25 gene (position -80 to -18; ref. 4) or κ -chain promoter from the V_H41 gene (*PvuII*-*Sfa*NI fragment, ref. 11) were labelled at either 5' end using a kinase reaction. End-labelled DNA fragments were partially methylated at purine residues by a modification of the method of Maxam and Gilbert²⁶ in which the dimethyl sulphate reaction was allowed to proceed for 7 min, then quenched with 1.5 M sodium acetate (pH 7.0), 1.0 M 2-mercaptoethanol and 100 μ g ml⁻¹ of poly-d(IC). The methylated DNA was precipitated twice with ethanol, washed with 70% ethanol, dried and resuspended in 10 mM Tris pH 8.0, 1 mM EDTA. A five-fold scaled-up binding reaction with the methylated DNA and WEHI-231 nuclear extract was subjected to native polyacrylamide gel electrophoresis (see Fig. 1 legend). After autoradiography of the wet gel, bands B1 and B3 and unbound DNA were cut out. The DNA from each band was electroeluted, bound to DEAE paper as described elsewhere²⁷, then eluted from the paper by heating to 68 °C for 1.5 h in elution buffer consisting of 20 mM Tris pH 8.0, 0.1 mM EDTA, 1 M NaCl. The supernatant was extracted twice with phenol and once with chloroform. After the addition of 1 μ g carrier transfer RNA, the DNA was precipitated twice with ethanol, rinsed with 70% ethanol, dried and redissolved in 100 μ l of 1.0 M piperidine. Reactions were incubated for 40 min at 90 °C and lyophilized twice. Equal amounts of radioactivity were subjected to electrophoresis through a 7% polyacrylamide/8 M urea sequencing gel followed by autoradiography at -70 °C with an intensifying screen. *c*, Wild-type or mutant octamer oligonucleotides were cloned into the *Bam*HI site of the PUC18 polylinker (T.W. *et al.*, manuscript in preparation). DNA fragments containing the wild-type or mutant octamers were excised by *Hind*III and *Eco*RI digestion of these plasmids and purified by polyacrylamide gel electrophoresis. Aliquots (25 ng) of these fragments were added as competitors to binding reactions of WEHI-231 nuclear extract with the radiolabelled κ -promoter fragment probe (~0.1 ng per reaction) and the products were analysed in the electrophoretic mobility shift assay (see Fig. 1 legend).

nuclear extract. DNA fragments in complexes B1 and B3 as well as the unbound fragments were isolated from a preparative native polyacrylamide gel, cleaved with piperidine and resolved on a sequencing gel (Fig. 3a). The complexed DNA from the two strands of both the H-chain promoter (Fig. 3a, panels 1, 2) and the κ promoter (panels 3, 4) was deficient in DNA modified at certain residues within the octamer motif, as indicated by arrows in Fig. 3a and summarized in Fig. 3b. Both guanine and, unexpectedly, adenine residues were cleaved in this experiment, although guanine residues reacted more efficiently. Most importantly, though, NF-A1 and NF-A2 had indistinguishable DNA binding sites by this analysis.

The fine specificity of the octamer binding proteins in WEHI-231 nuclear extract was analysed further by testing the ability of single transversion mutations in the octamer motif to compete with the wild-type octamer sequence for binding. Synthetic octamer oligonucleotides representing either the wild-type octamer sequence, ATTTGCAT, or transversion mutations at

positions 2-8 were cloned into the PUC18 polylinker. These DNA constructions also maintained a thymine residue two base pairs upstream of the initial adenine which is well conserved in immunoglobulin promoters^{1,3}. While the wild-type octamer-containing fragment competed away B1 and B3 equally well, the mutant octamer-containing fragments either did not compete at all or, in the case of the mutant ATTTGAAT, competed marginally (Fig. 3c). The competition by the wild-type octamer cloned into the PUC18 polylinker demonstrated directly that binding by the octamer-specific proteins does not require any additional sequences within the κ -promoter region. As expected from the high degree of conservation of the octamer sequence motif within immunoglobulin promoters *in vivo*, mutations at the seven positions tested interfered significantly with binding of octamer-specific proteins. Again, this fine-specificity analysis failed to reveal any differences in the DNA binding sites of NF-A1 and NF-A2.

Bacterial lipopolysaccharide (LPS) has pleiotropic effects on

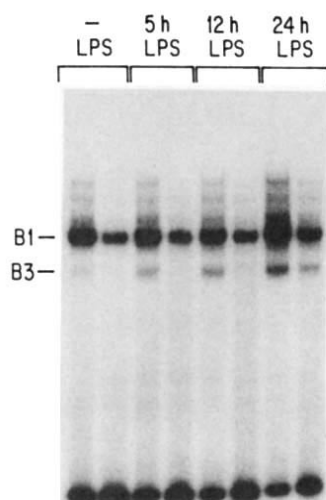


Fig. 4 Induction of the lymphoid-specific octamer binding protein by LPS. 70Z/3 cells were grown in roller bottles to $3-5 \times 10^5$ cells/ml in a 1-l volume and treated with LPS ($10 \mu\text{g ml}^{-1}$) for 5 h, 12 h or 24 h, after which nuclear extracts were prepared. These extracts as well as uninduced 70Z/3 nuclear extract were tested at 8 μg and 4 μg of protein in the electrophoretic mobility shift assay.

many B-lymphoid cell lines, including the ability to activate immunoglobulin gene transcription¹²⁻¹⁸, and we therefore tested its ability to alter the expression of octamer binding proteins. We prepared nuclear extracts from the LPS-responsive pre-B-cell line 70Z/3 (ref. 12) before and after treatment with LPS for 5, 12 or 24 hours. When equivalent amounts of nuclear extract protein were tested for octamer binding (Fig. 4), NF-A2 increased sixfold following 24 h treatment by LPS whereas NF-A1 was virtually unchanged. When another pre-B-cell line, PD (ref. 19), was treated for 24 h with LPS, again NF-A2 was induced fourfold by LPS without a corresponding increase in NF-A1 (data not shown). When cycloheximide was added 20 min before a 6.5-h LPS treatment of 70Z/3 cells, no induction of NF-A2 was observed (data not shown). Thus, the lymphoid-specific octamer binding protein NF-A2 is preferentially induced in lymphoid cell lines by treatment with LPS and this induction requires new protein synthesis. Because the induction of κ -chain transcription in 70Z/3 cells by LPS does not require new protein synthesis^{17,18}, an increased amount of NF-A2 is apparently not required for the induction of κ -chain transcription.

Recent experiments (T.W. *et al.*, manuscript in preparation) demonstrate that a synthetic promoter consisting of an octamer oligonucleotide placed upstream of a TATA box is sufficient for lymphoid-specific transcription. Thus, the octamer motif seems to be necessary^{2,4,6,8,9} and sufficient for lymphoid-specific promoter activity. In apparent contradiction, the human histone H2B gene promoter, which is presumably not tissue-specific, has an octamer positioned 17 base pairs upstream of a TATA box and mutation of the octamer significantly decreases transcription *in vitro*¹⁰. A possible explanation is that proteins bound to sequence motifs flanking the histone H2B octamer could be required to interact with NF-A1 in order to generate an active transcription complex. The H-chain promoter, which contains no other conserved sequence motifs besides the octamer, would then be inactive in non-lymphoid cells. NF-A2 might be able to stimulate transcription in the absence of other DNA binding factors. Thus, the lymphoid specificity of NF-A2 would explain the lymphoid specificity of immunoglobulin promoters.

Recent studies suggest that in lymphoid, but not non-lymphoid, cells the octamer motif contributes to H-chain enhancer activity^{20,21}. The H-chain enhancer has been shown to be equivalently active in pre-B, B- and myeloma cell lines²², all of which contain NF-A2. Thus, NF-A2 may be involved in the

lymphoid-specific activity of the H-chain enhancer as well as in promoter activity.

The discovery of distinct forms of a eukaryotic sequence-specific DNA binding protein raises the possibility that a single transcriptional control sequence is bound by different *trans*-acting factors in different tissues and under different conditions of cellular activation. The binding of distinct *trans*-acting factors to a single regulatory motif could influence the transcriptional activity of a gene in specific tissues.

This work was supported by a grant from the American Cancer Society (1M-355 Q) to D.B. and by grants from NIH (R01-GM32467, R01-CA38660), partially from an NCI Cancer Center Support (core) (P30-CA14051) and from NSF (PCM-8200309) to P.A.S. L.S. and H.S. are fellows of the Jane Coffin Childs Memorial Fund for Medical Research. R.S. is a fellow of the Damon Runyon-Walter Winchell Cancer Fund. T.W. is a fellow of the Deutsche Forschungsgemeinschaft.

Received 10 June; accepted 21 July 1986.

1. Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2650-2654 (1984).
2. Bergman, Y., Rice, D., Grosschedl, R. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7041-7045 (1984).
3. Falkner, F. G. & Zachau, H. G. *Nature* **310**, 71-74 (1984).
4. Grosschedl, R. & Baltimore, D. *Cell* **41**, 885-897 (1985).
5. Mason, J. O., Williams, G. T. & Neuberger, M. S. *Cell* **41**, 479-487 (1985).
6. Foster, J., Stafford, J. & Queen, C. *Nature* **315**, 423-425 (1985).
7. Queen, C. & Stafford, J. *Molec. cell. Biol.* **4**, 1042-1049 (1984).
8. Gopal, V. T., Shimada, T., Baur, A. W. & Nienhuis, A. W. *Science* **229**, 1102-1104 (1985).
9. Picard, D. & Schaffner, W. *EMBO J.* **4**, 2831-2838 (1985).
10. Sive, Hazel, L., Heintz, N. & Roeder, R. G. *Molec. cell. Biol.* (in the press).
11. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. *Nature* **319**, 154-158 (1986).
12. Paige, C. J., Kincade, P. W. & Ralph, P. J. *Immun.* **121**, 641-647 (1978).
13. Rosenberg, N., Siden, E. & Baltimore, D. in *B Lymphocytes in the Immune Response* (eds Cooper, M., Mosier, D., Scher, I. & Vitteta, E.) 379-386 (Elsevier, Amsterdam, 1979).
14. Perry, R. P. & Kelley, D. E. *Cell* **18**, 1333-1339 (1979).
15. Alt, F. A., Rosenberg, N., Casanova, R. J., Thomas, E. & Baltimore, D. *Nature* **296**, 325-331 (1982).
16. Lanier, L. L. *J. Immun.* **129**, 1130-1137 (1982).
17. Nelson, K. J., Kelley, D. E. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5305-5309 (1985).
18. Wall, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 295-298 (1986).
19. Lewis, S., Rosenberg, N., Alt, A. & Baltimore, D. *Cell* **30**, 807-816 (1982).
20. Waslylyk, C. & Waslylyk, B. *EMBO J.* **5**, 553-560 (1986).
21. Sen, R. & Baltimore, D. *Cell* (in the press).
22. Gerster, T., Picard, D. & Schaffner, W. *Cell* **45**, 45-52 (1986).
23. Pierce, J. H. & Aaronson, S. A. *J. exp. Med.* **156**, 873-887 (1982).
24. Braun, J. *J. Immun.* **130**, 2113-2116 (1983).
25. Dignam, J. D., Lebowitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
26. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 449-525 (1980).
27. Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. *Analyt. Biochem.* **112**, 295-298 (1981).

A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma

Stephen H. Friend*†, Rene Bernards*, Snezna Rogelj*, Robert A. Weinberg*‡, Joyce M. Rapaport§, Daniel M. Albert§ & Thaddeus P. Dryja§

* Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA

† Division of Hematology-Oncology, The Children's Hospital, Dana-Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

§ Department of Ophthalmology, Harvard Medical School and Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114, USA

The genomes of various tumour cells contain mutant oncogenes that act dominantly, in that their effects can be observed when they are introduced into non-malignant cells¹⁻⁴. There is evidence for another class of oncogenes, in which tumour-predisposing mutations are recessive to wild-type alleles⁵⁻⁷. Retinoblastoma is a prototype biological model for the study of such recessive

oncogenes⁸. This malignant tumour, which arises in the eyes of children, can be explained as the result of two distinct genetic changes, each causing loss of function of one of the two homologous copies at a single genetic locus, *Rb* (refs 9–12), assigned to the q14 band of human chromosome 13 (refs 13–22). Mutations affecting this locus may be inherited from a parent, may arise during gametogenesis or may occur somatically. Those who inherit a mutant allele at this locus have a high incidence of non-ocular, second tumours²³, almost half of which are osteosarcomas believed to be caused by the same mutation^{24,25}. Here we describe the isolation of a complementary DNA segment that detects a chromosomal segment having the properties of the gene at this locus. The gene is expressed in many tumour types, but no RNA transcript has been found in retinoblastomas and osteosarcomas. The cDNA fragment detects a locus spanning at least 70 kilobases (kb) in human chromosome band 13q14, all or part of which is frequently deleted in retinoblastomas and osteosarcomas.

A previous report showed that a 1.5-kb human DNA sequence termed H3-8, selected from a human chromosome 13 lambda phage library²⁶, could detect deletions involving 13q14 in 3 out of 37 retinoblastomas²⁷. Based on the presumed proximity of the H3-8-homologous sequence to the *Rb* locus, we used chromosome walking techniques to isolate and map 30 kb of surrounding genomic DNA²⁷. DNA derived from human-mouse hybrid cell lines confirms the assignment of selected single-copy fragments generated during the chromosome walk to chromosome 13.

One of these single-copy fragments, p7H30.7R, recognizes a DNA sequence in the mouse genome and in human chromosome 13. The conservation of this DNA sequence between mice and humans suggested that the cloned fragment contains a coding exon of a gene. We therefore tested the ability of this fragment to hybridize to RNA derived from retinoblastoma cells and adenovirus-12-immortalized human retinal cells²⁸, and found that it recognizes a 4.7-kb RNA transcript in the retinal cell line that is not detectable in four retinoblastomas (data not shown), consistent with its association with the *Rb* gene.

We next constructed a cDNA library from RNA of the human retinal line using the bacteriophage cloning vector λ gt-11 (ref. 29) and screened this library with the p7H30.7R probe. Several cDNA clones were isolated with similar physical maps. The longest of these, carrying a 4.7-kb insert, was named p4.7R (Fig. 1) and used to screen cytoplasmic RNA prepared from the adenovirus-12-immortalized retinal cell line, from four retinoblastoma cell lines and from one osteosarcoma cell line. The cDNA probe detects a transcript present in the adenovirus-immortalized retinal cells but not detectable in the tumour cell lines (Fig. 2). Northern blot analysis of poly(A) messenger RNA from adenovirus-immortalized retinal cells yields identical results (data not shown). A 4.7-kb transcript hybridizing to the probe is also detected in other tumour types, including small-cell carcinoma of the lung, neuroblastoma, renal carcinoma and melanoma (Fig. 3), as well as normal adult human retina, spleen and liver (data not shown).

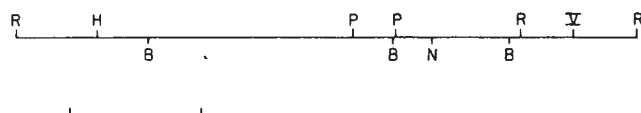


Fig. 1 Physical map of the p4.7R cDNA insert. Restriction enzyme sites are indicated as follows: R, *EcoRI*; H, *HindIII*; B, *BglII*; P, *PstI*; N, *NcoI*; V, *EcoRV*.

Methods. A cDNA library was constructed in λ gt11 from adenovirus-immortalized retinal cells²⁸ using the technique of Gubler and Hoffman²⁹. This technique involved oligo(dT) priming and second-strand synthesis with DNA polymerase I and RNase H. The double-stranded cDNA was methylated at internal *EcoRI* sites, ligated to *EcoRI* linkers, and size-selected for cDNA fragments >2.3 kb by agarose gel electrophoresis. Bar, 1 kb.

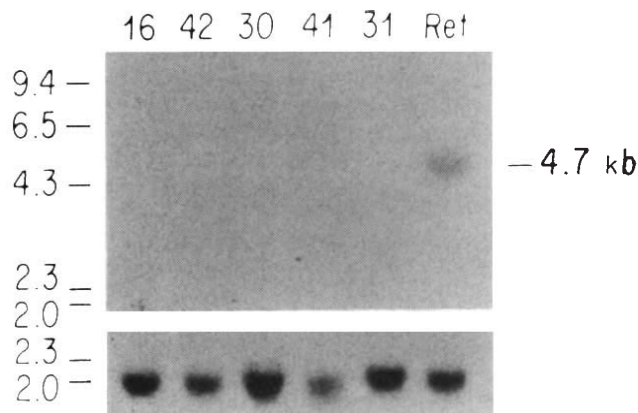


Fig. 2 Northern blot analysis of RNA isolated from retinal and osteosarcoma cells. RNA (20 μ g) was loaded from osteosarcoma tumour No. 16; four retinoblastomas: Nos 42, 30 (Y-79, ref. 30), 41 (WERI-1, ref. 31) and 31; and the adenovirus-12-transformed retinal cells (ref. 28). *a*, Samples probed with p4.7R insert. The transcript detected in the retinal cells is 4.7 kb in size. Molecular mass markers indicated on the left. *b*, After washing the filter, it was rehybridized with a probe derived from the rat tubulin locus³². **Methods.** Total cytoplasmic RNA was isolated using the Nonidet-P-40 extraction technique, resolved by electrophoresis through a 1% agarose-formaldehyde gel, and transferred onto nitrocellulose as described by Schrier *et al.*³³. Radiolabelling of probes was performed using the random-primer method³⁴.

The ubiquitous presence of this transcript contrasts to its apparent absence in the retinoblastoma and osteosarcoma specimens. Indeed, its absence in retinoblastomas and osteosarcomas might be central to the pathogenesis of these two tumour types. Specifically, our data are consistent with a close association between the DNA template of the 4.7-kb transcript and the *Rb* gene. Further support of this association requires analysis of the tumour DNA sequences recognized by the p4.7R probe.

The probe allowed a survey of a 70-kb chromosomal region. Preliminary mapping was performed by ordering the genomic *HindIII* restriction endonuclease fragments present in normal human DNAs that are reactive with this probe (Fig. 5). Because this map was generated by use of a cDNA probe, it includes only genomic fragments that specify sequences found in exons. We made similar maps of the fragments generated by *TaqI* and *XbaI* (data not shown). All the DNA fragments included in this map and used in the following studies were assigned to chromosome 13 (data not shown).

The p0.9R probe, the subclone containing the 0.9 kb *EcoRI*/*EcoRI* fragment deriving from the right end of p4.7R, reacts with three *HindIII* fragments (1.2, 1.5 and 5.8 kb) that occur in genomic lambda phage clones of this region²⁷ and are included in Fig. 5. This probe reacts with a low copy-number repeat sequence present in the human genome (Fig. 4b). The following analyses were possible despite the detection of these repeated sequences.

DNA was isolated from a set of tumours from unrelated patients, each of which was thought to arise because of inactivating mutations at the *Rb* locus. The set consisted of 40 retinoblastomas (including tumours from patients with sporadic and familial disease), 8 osteosarcomas and 2 undifferentiated tumours of unknown cellular origin arising in patients with hereditary retinoblastoma.

Screening of tumour DNA samples using Southern blot analysis and probes derived from p4.7R reveals three types of deviant restriction-fragment patterns: totally absent fragments, reflecting apparent homozygous deletions; under-represented fragments, reflecting apparent heterozygous deletions; and fragments of altered size, reflecting partial deletion, sequence rearrangement

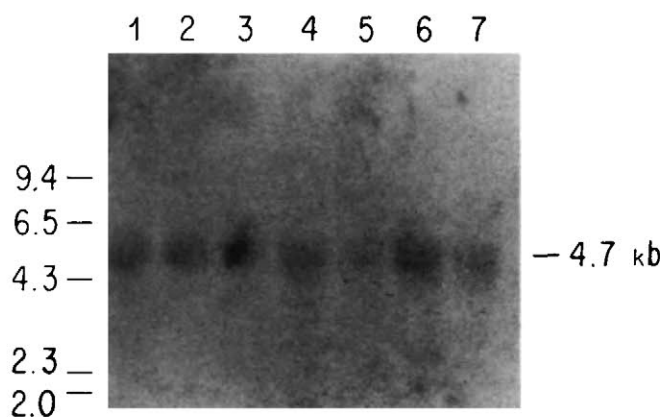


Fig. 3 Northern blot analysis of RNA isolated from various tumour cell lines. RNA was prepared and loaded as in Fig. 2. Lane 1, LX-1 (small-cell carcinoma of the lung); lane 2, SKNSH (neuroblastoma); lane 3, CAKI-1 (renal clear cell carcinoma); lane 4, MEL-5 (melanoma); lane 5, SKNMC (neuroblastoma); lane 6, SKRC-2 (renal carcinoma); lane 7, ACHN (renal adenocarcinoma). 32 P-labelled p4.7R insert was used as a probe. Positions of molecular mass markers are indicated on the left.

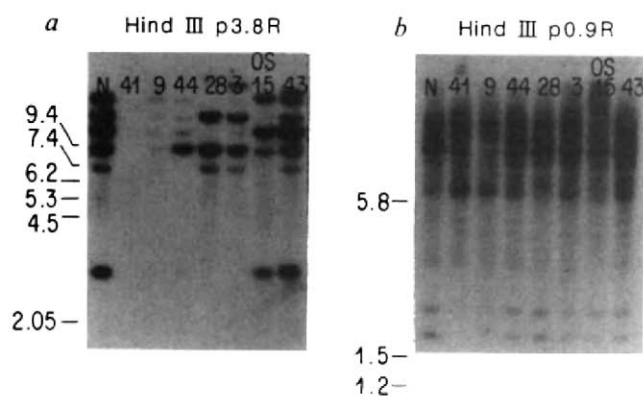


Fig. 4 Southern blot analysis of retinoblastoma, osteosarcoma and normal DNAs probed with subclones of p4.7R (Fig. 1). The first lane (N) contains lymphocyte DNA from a normal, healthy human. The adjacent lanes have DNA from the following tumours: retinoblastoma No. 41 (WERI-1, ref. 31); retinoblastoma No. 9; retinoblastoma No. 44; retinoblastoma No. 28; retinoblastoma No. 3; osteosarcoma OS-15 (ref. 24); and retinoblastoma No. 43. DNA (4 µg per lane) was digested with *Hind*III. *a*, Probed with a 3.8-kb *Eco*RI/*Eco*RI subcloned fragment of p4.7R (see Fig. 1). *b*, Probed with a 0.9-kb *Eco*RI/*Eco*RI subcloned fragment of p4.7R (Fig. 1). **Methods.** DNA isolation, restriction endonuclease digestion of DNA samples, agarose gel electrophoresis, Southern blotting and hybridization were all performed according to standard methods. Radiolabelling was performed using the random-primer method³⁴.

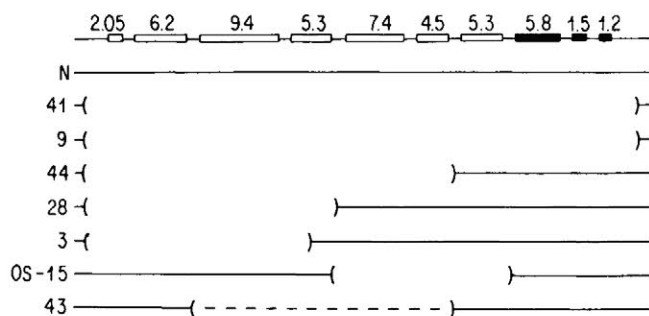


Fig. 5 Schematic representation of deletions in the genomic DNA from six retinoblastomas and an osteosarcoma. The genomic *Hind*III fragments detected by the 3.8- and 0.9-kb *Eco*RI fragments, denoted by open and closed bars, respectively (see Fig. 4), are shown at the top according to their relative positions in the genome and labelled with sizes in kilobases. Normal and tumour DNAs, numbered as in Fig. 4, have solid lines where both copies of a given genomic fragment are present. Dotted line between brackets, heterozygous deletion; open brackets, homozygous deletion.

or alteration of a restriction endonuclease recognition site. At least 30% of the tumour DNAs have one of these abnormalities. In contrast, analysis of leukocyte DNA from 18 normal individuals shows a uniform pattern of restriction fragments, hence the abnormal patterns of fragments probably do not represent frequently occurring, incidental polymorphisms.

Homozygous deletion of restriction fragments in six of the tumours are shown in Fig. 4. DNA from tumours passaged in tissue culture (for example, No. 41) or in nude mice (for example, OS-15) lack certain restriction fragments normally detected by the probe. DNA isolated directly from surgical specimens (for example, Nos 9, 44 and 3) has a low concentration of certain normal restriction fragments, amounting to <10% the normal diploid dose. These faint bands result from small numbers of normal cells from adjacent tissue contaminating tumour cells with homozygous deletions. In some tumours (Nos 28, 3 and 43) deletions also result in the appearance of *Hind*III fragments of novel length.

The extent of the deletions present in these tumours is mapped in Fig. 5, and is similar to results obtained with *Taq*I or *Xba*I (data not shown). Most of the deletions, including those present in tumours 41, 9, 44, 28 and 3, have an endpoint in the transcription unit defined by the probe and extend apparently leftwards beyond this unit (Fig. 5). These deletions may have signified that a gene to the left of this transcription unit was the target of the oncogenic mutations. However, we have found two instances in which the deletions involve only internal regions of this unit, leaving adjacent DNA unaffected.

The first of these two cases concerns retinoblastoma tumour No. 43, which carries a deletion involving DNA sequences that map entirely within the confines of the transcription unit defined by the probe (Figs 4, 5). This deletion is present in a heterozygous state. Thus, Fig. 4 shows that two restriction fragments created by *Hind*III cleavage (of 2.05 and 6.2 kb) and which map to the left part of this genomic region are present in normal, diploid amounts. Similarly, fragments found at the right end (*Hind*III: 1.5, 1.2 and 5.8 kb) are also found in normal dosage. In contrast, fragments from the centre of this region (*Hind*III: 9.4, 5.3, 7.4 and 4.5 kb) are present in about half the normal amounts. A novel *Hind*III fragment of 10.0 kb arises from the partial deletion of one copy of the normally present 9.4-kb *Hind*III fragment.

The second instance of an internal deletion involves an osteosarcoma, OS-15, which arose in a patient who had no retinoblastoma or family history of retinoblastoma. The extent

of the deletion, which was homozygous in this case, is indicated in Figs 4 and 5 by the absence of the 7.4- and 4.5-kb *Hind*III fragments. The internal deletions in tumours 43 and OS-15 suggest that the chromosomal segment recognized by p4.7R is the common target for these mutations.

Note that the genetic origin of the deletions found in 3 of the 50 tumours has been shown previously to be somatic in two cases and germinal in the third. Therefore, these deletions are not inherited polymorphisms²⁷. These data, taken together with the internal deletions described above, strongly suggest that this segment is the carrier of the DNA sequences of the *Rb* gene.

Earlier work shows the absence of a detectable transcript in five tumours from our set (Fig. 2). The apparent absence of transcript from these tumours is presumably important in their aetiology and lesions affecting the DNA of this gene must be frequently involved in the inactivation of expression. Accordingly, we examined the DNA of the four retinoblastomas that

lacked detectable transcript. One of these has a homozygous deletion of the region detected by p4.7R (tumour 41, Fig. 4). However, the other three retinoblastomas have no abnormality. We deduce that tumours having no grossly aberrant genomic fragments may nevertheless have mutations affecting this region. A precedent for the role of small structural changes as a cause of gene inactivation is provided by studies of other human disease loci in which gross changes in DNA structure represent the minority of mutations responsible for loss of function³⁵.

Fifteen years ago, Knudson provided a theoretical basis for retinoblastoma tumorigenesis by suggesting that minimally two genetic events are required to trigger tumour development⁸. These events may represent loss-of-function mutations at the retinoblastoma locus, and other tumours may arise by a similar mechanism via recessive mutations at other loci³⁶. Evidence in support of this theory comes from the study of a diverse group of tumours, including Wilms tumour, hepatoblastoma, rhabdomyosarcoma³⁷, uveal melanoma³⁸ and bladder cell carcinoma³⁹. A recessive oncogene from the fruitfly with tumour-inhibiting properties has recently been isolated⁴⁰.

We have isolated a human cDNA sequence apparently representing one of this class of genes. This cDNA detects a chromosomal segment having properties of the *Rb* gene which maps to human chromosome 13q14. Part or all of the sequence is deleted in some retinoblastomas and osteosarcomas, and some of the deletions affect only an internal region of the segment.

We next seek to determine whether some malignant properties of retinoblastoma and osteosarcoma cells can be reverted on introduction of cloned sequences thought to represent the *Rb* gene. The isolation of the p4.7R clone should also provide a powerful diagnostic reagent that will help to determine whether a particular retinoblastoma stems from an inherited genetic lesion or from a new mutation of somatic or germinal origin. Furthermore, the variable expression of this gene in different tumours may provide a new method by which tumours are classified.

We thank Sang-Ho Park, John E. Price, Frederick P. Li, Raju Chaganti, Jennifer M. Joyce and Theresa Dryden for technical assistance; David Page, Michael Gilman, Cori Bargmann and David Stern for helpful comments; R. T. M. J. Vaessen and A. J. van der Eb for the gift of the adenovirus-transformed retinal cells; Dr Samuel Lett in the Department of Genetics for efforts in developing the H3-8 probe; Robert Petersen, David Walton, Johan Zwaan, Charles Regan, Ted Sery, Webster Cavenee, William Benedict, Brenda Gallie, Raju Chaganti, Allen Goorin, David Abramson, James Epstein, Steven Blattner and Z. Nicholas Zakov for help in obtaining tumour specimens. R.B. is a fellow of The Netherlands Organisation for the Advancement of Pure Research; S.R. is an NIH fellow; R.A.W. is an American Cancer Society Research Professor. Supported by NIH grants EY05321 and CA39826, a grant from the American Business Foundation for Cancer Research, and by gifts from the Wytres, Friedman, Smith and Gorfinkle families to the Taylor R. Smith Laboratory of the Massachusetts Eye and Ear Infirmary.

Received 24 July; accepted 19 August 1986.

1. Bishop, J. M. *Cell* **42**, 23-28 (1985).
2. Land, H., Parada, L. & Weinberg, R. A. *Science* **222**, 771-778 (1983).
3. Heldin, C. H. & Westermark, B. *Cell* **37**, 9-20 (1984).
4. Sporn, M. B. & Roberts, A. B. *Nature* **313**, 745-747 (1980).
5. Klein, G. & Klein, E. *Nature* **315**, 190-195 (1985).
6. Green, A. R. & Wyke, J. A. *Lancet* **ii**, 475-477 (1985).
7. Knudson, A. G. Jr *Cancer Res.* **45**, 1437-1443 (1985).
8. Knudson, A. G. Jr *Proc. natn. Acad. Sci. U.S.A.* **68**, 820-823 (1971).
9. Murphree, A. L. & Benedict, W. F. *Science* **223**, 1028-1033 (1984).
10. Godbout, R. et al. *Nature* **304**, 451-453 (1983).
11. Cavenee, W. K. et al. *Nature* **305**, 779-784 (1983).
12. Dryja, T. P. et al. *N. Engl. J. Med.* **310**, 550-553 (1984).
13. Sparkes, R. S. et al. *Science* **208**, 1042-1044 (1980).
14. Yunis, J. J. & Ramsey, N. Am. J. *Dis. Child.* **132**, 161-163 (1978).
15. Vogel, F. *Hum. Genet.* **52**, 1-54 (1979).
16. Motegi, T. *Hum. Genet.* **61**, 95-97 (1982).
17. Balaban, G. et al. *Cancer Genet. Cytogenet.* **6**, 213-221 (1982).

18. Chaum, E. et al. *Cytogenet. Cell Genet.* **38**, 82-91 (1984).
19. Sparkes, R. S. et al. *Science* **219**, 971-973 (1983).
20. Connolly, M. J. et al. *Hum. Genet.* **65**, 122-124 (1983).
21. Mukai, S. et al. *Am. J. Ophthalmol.* **97**, 681-685 (1984).
22. Halloran, S. B. et al. *Arch. Ophthalmol.* **103**, 1329-1331 (1985).
23. Abramson, D. H. et al. *Ophthalmology* **91**, 1351-1355 (1984).
24. Dryja, T. P. et al. *Am. J. hum. Genet.* **38**, 59-66 (1986).
25. Hansen, M. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6216-6220 (1985).
26. Lalonde, M. et al. *Cancer Genet. Cytogenet.* **13**, 283-295 (1984).
27. Dryja, T. P. et al. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
28. Vaessen, R. T. M. J. et al. *EMBO J.* **5**, 335-341 (1986).
29. Gubler, U. & Hoffman, B. *Gene* **25**, 263-269 (1983).
30. Reid, T. W. et al. *J. natn. Cancer Inst.* **53**, 347-360 (1974).
31. McFall, R. C., Sery, T. W. & Makadon, M. *Cancer Res.* **37**, 1003-1010 (1977).
32. Lemischka, I. R. et al. *J. molec. Biol.* **150**, 101-120 (1981).
33. Schrier, P. I. et al. *Nature* **305**, 771-775 (1983).
34. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
35. Kunkel, L. M. et al. *Nature* **322**, 73-77 (1986).
36. Comings, D. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3324-3328 (1973).
37. Koufos, A. et al. *Nature* **316**, 330-334 (1985).
38. Mukai, S. & Dryja, T. P. *Cancer Genet. Cytogenet.* **22**, 45-53 (1986).
39. Fearon, E. R. et al. *Nature* **318**, 377-380 (1985).
40. Mechler, B. M., McGinnis, W. & Gehring, W. J. *EMBO J.* **4**, 1551-1557 (1985).

Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene

Anthony P. Monaco*†, Rachael L. Neve*†, Chris Colletti-Feener*, Corlee J. Bertelson*, David M. Kurnit* & Louis M. Kunkel*†‡

* Division of Genetics, Mental Retardation Program, Department of Pediatrics, Harvard Medical School, The Children's Hospital, Boston, Massachusetts 02115, USA

† The Program in Neuroscience, Harvard University, Cambridge, Massachusetts 02138, USA

Duchenne muscular dystrophy (DMD) and the less severe Becker muscular dystrophy (BMD) are human X-linked muscle-wasting disorders that have been localized to the band Xp21 by genetic linkage analysis¹⁻⁹ and cytologically detectable abnormalities¹⁰⁻¹². A cloned DNA segment, DXS164 (or pERT87), has been shown to detect deletions in the DNA of unrelated DMD and BMD males¹³⁻¹⁵. Here we present the nucleotide sequence of two highly conserved DNA fragments from the DXS164 locus and their homologous sequences from the mouse X chromosome. One of the human conserved segments hybridized to a large transcript in RNA isolated from human fetal skeletal muscle and was used to isolate cDNA clones which cover approximately 10% of this transcript. The cDNA clones map to Xp21 and hybridize with a minimum of eight small regions that span 130 kilobases (kb) of the DXS164 locus. These expressed sequences are candidates for portions of the gene responsible for both DMD and BMD.

The cloned DNA of the DXS164 locus was expanded from 137 kb¹⁵ to 220 kb of contiguous DNA in a total of nine bidirectional chromosome walks made in human partial digest phage libraries. A schematic restriction map for the enzymes *EcoRI*, *HindIII*, and *KpnI* is presented in Fig. 1 for the entire length. Given that the DXS164 locus was found to be deleted in patients with DMD and BMD^{14,15} and was centrally located with respect to mutations which give rise to the disease¹⁵, we searched for regions that might be transcribed. As an alternative approach to Northern blot analysis of each individual DNA subclone from the DXS164 locus, a systematic search was undertaken for nucleotide sequence conservation across species. Approximately 50 unique subclones from the DXS164 locus were used individually as hybridization probes on Southern blots of *HindIII* digested DNA isolated from cebus, bovine, mouse, hamster and chicken. Most unique subclones were found to

‡ To whom correspondence should be addressed.

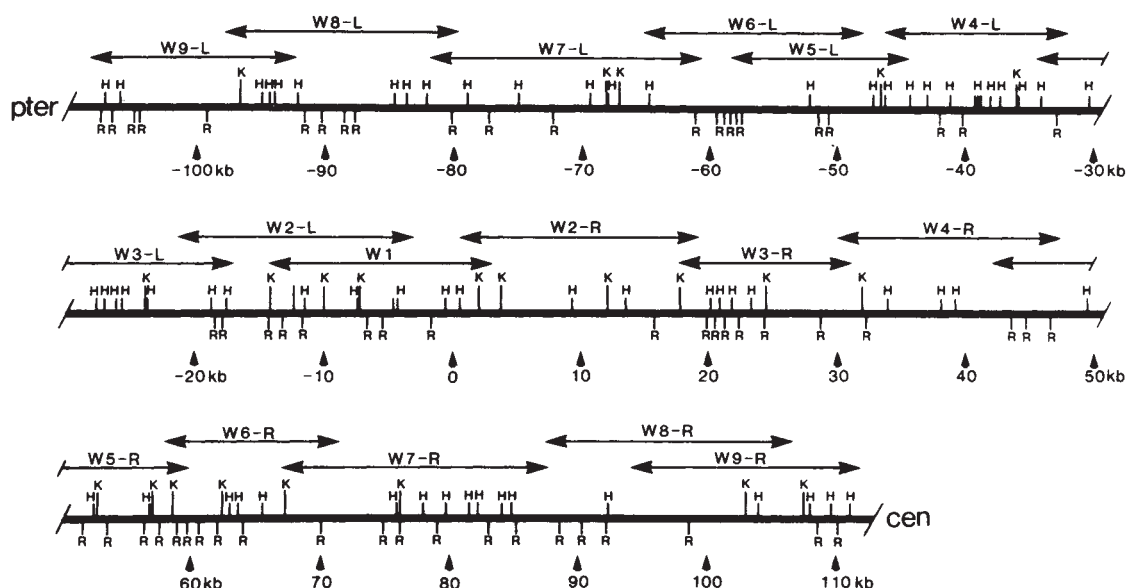
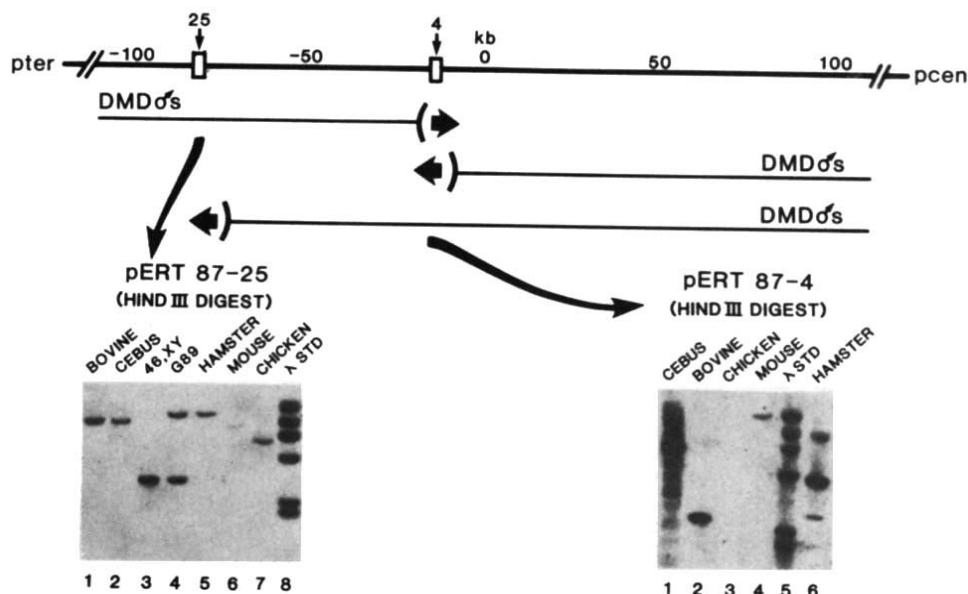


Fig. 1 Schematic restriction map of the DXS164 (pERT87) locus. The heavy line indicates the points of cleavage for the restriction enzymes *EcoRI* (R), *HindIII* (H) and *KpnI* (K). The coordinates show the distance in kilobases starting at zero for the position of the original pERT87 clone¹³⁻¹⁵. The terminus of the short arm of the X-chromosome is indicated on the left-hand side of the map and the centromere on the right as positioned by deletion analysis¹⁵. Lines with arrowheads represent overlapping cloned DNA segments isolated from human recombinant phage libraries¹⁴. Numbers on these lines detail the walk number and direction (left or right from zero). Phage clones with human inserts were plaque purified and DNA was isolated²⁸ and restriction mapped. Phage DNA digests were subcloned into the plasmid vectors pBR322²⁹, pUC18³⁰ and Blue-Scribe (Stratagene). Subclones were restriction mapped further with the enzymes *XbaI*, *BglII*, *PstI* and *BamHI*, to identify small fragments free of human repetitive elements. All unique sequence subclones were tested for human X-chromosome localization as previously described¹⁴. Methods and conditions described elsewhere^{2,14}.

Fig. 2 Cross-species homology of two subclones from the DXS164 locus. Dark black line, the 220 kb of cloned DNA of the DXS164 locus. Markers for every 50 kb are positioned across the map as well as two conserved subclones, pERT87-25 and pERT87-4, which are 70 kb apart. Two areas defined by deletion breakpoint analysis that may be important to the expression of the DMD gene¹⁵ are indicated below the schematic of the DNA map. Autoradiograph lower left, hybridization analysis of subclone pERT87-25 (a 1.95 *EcoRI*/*HindIII* fragment) to specific *HindIII* restriction fragments in 1.5 µg of the following DNA samples: bovine DNA (lane 1), cebus monkey DNA (2, provided by A. Scott), human male DNA (3), human X-chromosome DNA in a hamster DNA background (4, G89, ref. 32), hamster DNA alone (5, provided by J. Kang), mouse DNA (6, RAG cell line provided by G. Bruns) and chicken DNA (7). Radiolabelled *HindIII* digested λ DNA (8) was used as a size marker (2.0, 2.3, 4.4, 6.6, 9.4 and 23.5 kb). Autoradiograph lower right, hybridization of subclone pERT87-4 (2.3 kb of pooled *XbaI* fragments) to several *HindIII* fragments in cebus monkey DNA (1), and a single *HindIII* fragment in bovine DNA (2), mouse DNA (4) and a human X-chromosome band at 4.2 kb on a hamster DNA background (9.0 and 2.7 kb bands) (G89, lane 6). pERT87-4 did not hybridize to chicken DNA (3). Radiolabelled *HindIII* digested λ DNA is shown (5). Both Southern blots were done at high stringency (hybridize in 50% formamide, 4×SSC at 45°C; wash at 55°C for 1 h in 0.1% SDS and 0.1×SSC). Negative hybridization results with both pERT87-25 and pERT87-4 for DNA isolated from *Drosophila melanogaster* (provided by W. Bender), *Caenorhabditis elegans* (provided by S. Wood) and lobster (provided by H. Potter) are not shown.



hybridize to cebus primate DNA, and several hybridized weakly to bovine DNA. Only two subclones, pERT87-4 and pERT87-25, exhibited hybridization to all mammalian DNA samples tested (Fig. 2). Subclone pERT87-4 is from a 4.2 kb *HindIII* fragment located at -10 kb in the DXS164 locus, and hybridizes to specific *HindIII* fragments in mammalian but not chicken DNA. This subclone lies inside the 10 kb region of deletion overlap found

in most deletions analysed at the DXS164 locus¹⁵. Subclone pERT87-25 is from a 1.95 kb *EcoRI*/*HindIII* fragment located at -82 kb in the DXS164 locus, 70 kb toward the terminus from pERT87-4. It hybridizes to specific *HindIII* fragments in DNA from all mammals tested and strongly to chicken DNA. Subclone pERT87-25 is deleted from the DNA of two DMD males whose deletion breakpoints occur further toward the terminus¹⁵.

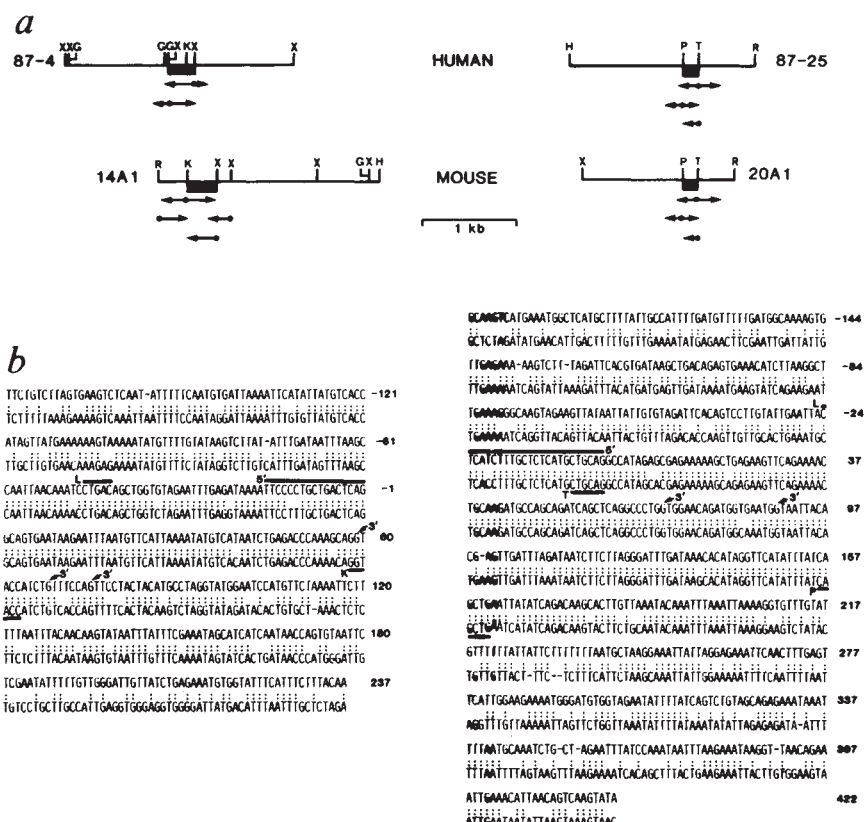


Fig. 4 Hybridization of three DNA segments to identical Northern blots³⁴ containing fetal skeletal muscle RNA. *a*, Hybridization of the conserved DNA fragment, pERT87-25, to total RNA (10.0 µg) isolated from cultured human myoblast cells (lane 1), fetal skeletal muscle (2), HeLa transformed cells (3) and poly(A)⁺ selected RNA (2.0 µg) isolated from a separate fetal skeletal muscle (4). Left, radiolabelled *Hind*III digested λ DNA as size markers. *b*, Hybridization of a cloned segment of the PGK gene at 2.0 kb to the same Northern blot as in *a*. *c*, Hybridization of cDNA 5' to the same Northern blot as in *a*.

Methods. Total cellular RNA was isolated from cells and tissues using 5.0 M guanidinium isothiocyanate³⁵. Total and polyadenylated RNA³⁶ were separated by electrophoresis on 1.0% formaldehyde agarose gels and transferred to Pall Biotryne A membrane following the protocol recommended by the manufacturer. Hybridizations of radiolabelled probes (specific activities 2–5 × 10⁸ d.p.m. per µg) were done in 50% formamide and 5 × SSC at 45 °C and washed in 1.0 × SSC and 0.1% SDS at 55 °C for 1 h. All autoradiographs presented were done at –70 °C with DuPont intensifying screens for four days. Between hybridizations the membranes were stripped by washing in 50% formamide and 0.01 M sodium phosphate buffer (pH 7.0) for one hour at 65 °C. Human fetal tissue was obtained from 18–22 week-old abortuses under approved protocols at the Brigham and Women's Hospital. Human cultured myoblast cells were provided by Dr B. Brown. Human B-cell RNA was provided by Dr S. Orkin and T-cell RNA by Dr B. Rover-Pokora.

isolated from a hamster-mouse somatic cell-hybrid bearing a mouse X chromosome (data not shown). A comparison of the restriction maps with several enzymes for each of the human and mouse subclones is presented in Fig. 3a. The pERT87-25 human and mouse subclones were each shown to contain a cross-hybridizing 150 bp region located between conserved *Pst*I and *Pvu*II sites. The pERT87-4 human and mouse subclones share a conserved *Kpn*I site around an approximately 260 bp

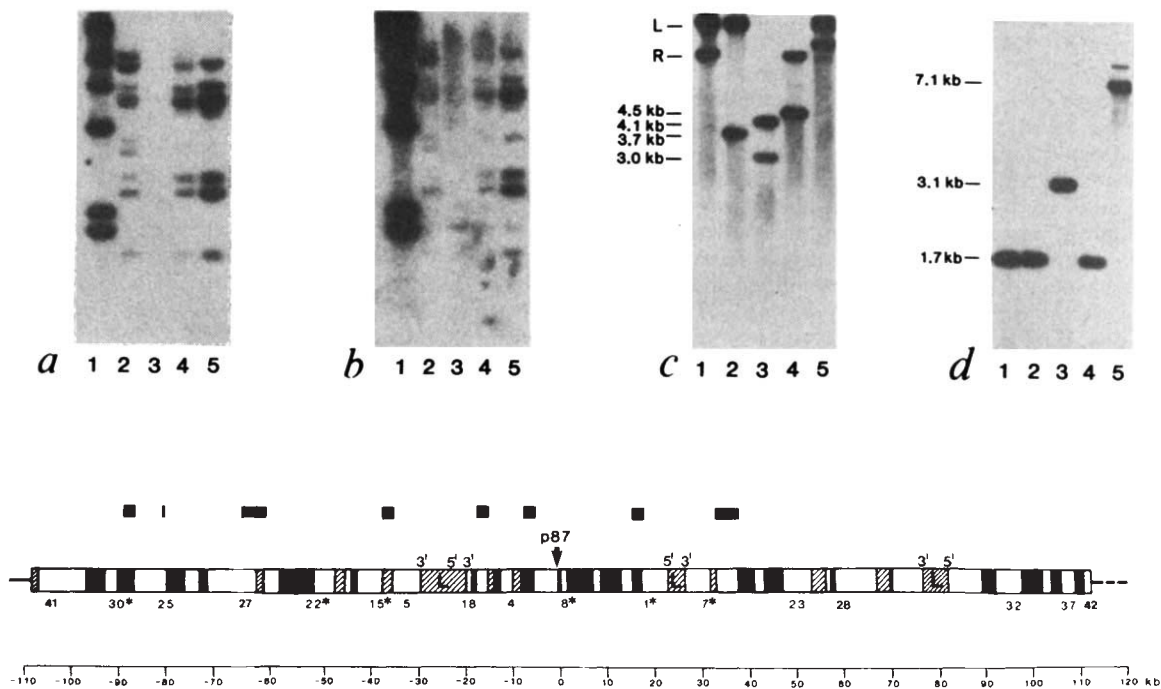


Fig. 5 Hybridization results for cDNA clones 5 and 6 to genomic Southern blots and to cloned phage DNA from the DXS164 locus. *a*, Southern blot analysis of cDNA clone 5 to *Hind*III digested DNA samples (1.0 μ g per lane) to indicate Xp21 localization. The cDNA clone 5 detects eight different *Hind*III fragments (~1.7, 2.7, 3.0, 5.0, 5.5, 6.0, 6.6 and 8.0 kb) which all hybridize more intensely to 49,XXXXY DNA (lane 5) than to DNA with a single copy of the X chromosome short arm (lane 4), and are all missing in DNA isolated from a lymphoblastoid cell-line established from the male patient, B.B., who had a visible Xp21 deletion and three X-linked diseases¹¹ (lane 3). The same *Hind*III fragments are detected in an isolated human X-chromosome in a hamster DNA background (G89, lane 2) as well as several other restriction fragments, probably conserved hamster fragments. Radiolabelled *Hind*III digested λ DNA is given (lane 1). *b*, The cDNA clone 6 hybridizes to the same eight *Hind*III fragments as cDNA clone 5 plus two additional fragments, one at 4.2 kb which localizes to Xp21 and one at 2.2 kb which does not demonstrate X-chromosomal localization. *c*, The hybridization of cDNA clone 5 to two independent restriction fragments in the following enzyme digests of cloned DNA of walk 7 left of the DXS164 locus: lane 1, *Hind*III; 2, *Hind*III/*Eco*RI; 3, *Hind*III/*Sal*I; 4, *Hind*III/*Kpn*I; 5, *Eco*RI; L, left phage arm, R, right phage arm. *d*, The hybridization pattern of cDNA clone 5 to a potential exon in the following cloned phage DNA digests from walk 4 left: lane 1, *Hind*III; 2, *Hind*III/*Kpn*I; 3, *Eco*RI/*Kpn*I; 4, *Eco*RI/*Hind*III; 5, *Eco*RI. Bottom, schematic representation of the DXS164 locus, dark boxes above indicate the location of the smallest restriction fragments which hybridize with the cDNA clones (examples of hybridization in Fig. 5c and d). The schematic map of the DXS164 locus is subdivided into highly repeated regions indicated by dark boxes, moderately repeated regions (a unique Xp21 band among many other hybridizing fragments) by hatched boxes, unique subclones by white boxes and the LINE sequence³⁷ by L. The numerical designations of unique subclones are indicated below their corresponding map position and those that recognize RFLPs are marked with an asterisk. The LINE sequences were further characterized as to their length and orientation using cloned 5' and 3' ends of a human LINE sequence as hybridization probes (provided by B. Schmeckpeper) as indicated above the map. The hybridization pattern of the two cDNA clones to phage DNA from the DXS164 locus suggests that two introns contain copies of LINE sequence homology.

human *Xba*I fragment that cross-hybridizes to a 270 bp mouse *Kpn*I/*Xba*I fragment.

The nucleotide sequence of the human and mouse subclones was determined to identify potential open reading frames coding for uninterrupted amino acid stretches on either strand through the conserved regions (sequencing strategy given in Fig. 3a). Once open reading frames were found, we looked for homology to the consensus sequences for the 5' splice acceptor¹⁷, the 3' splice donor¹⁷ and the lariat acceptor^{18,19}. The only two potential open reading frames that were surrounded by nucleotide homology to all three consensus sequences are indicated in Fig. 3b with the human nucleotide sequence aligned above the mouse sequence for each subclone. Both pERT87-25 and pERT87-4 have their hypothetical 5' and 3' splice sequences oriented in the same direction toward the terminus of the human X chromosome and predict an RNA transcript that would initiate on the centromere side of the DXS164 locus and end towards the terminus of the short arm.

The two human conserved DNA fragments, pERT87-25 and pERT87-4, were used as hybridization probes to screen RNA isolated from human fetal tissues. pERT87-25 was found to

hybridize to two different human fetal skeletal muscle RNA samples (Fig. 4a). In total RNA isolated from one fetal skeletal muscle sample (lane 2) and poly(A)⁺ selected RNA isolated from a different fetal skeletal muscle sample (lane 4), pERT87-25 hybridizes to a large, discrete band at approximately 16 kb with a smear of lighter hybridization below. This hybridization pattern is similar to that seen by the apolipoprotein B cDNA to its large 19 kb transcript in RNA isolated from human small intestine²⁰. pERT87-25 failed to hybridize to total RNA isolated from cultured human myoblast cells (lane 1) and HeLa-transformed cells (lane 3). When pERT87-4 was used to probe this same Northern blot, it failed to detect the large fetal muscle transcript seen by pERT87-25 and was also negative in myoblast and HeLa RNA (data not shown). Perhaps pERT87-4 is used at low frequency as an alternatively spliced exon in muscle RNA²¹, below the level of detection on Northern blots, since the conserved nucleotide sequence in both human and mouse indicates that it may be an exon. All four RNA samples exhibited strong hybridization at 2.0 kb with a cloned segment of the constitutively expressed X-linked phosphoglycerate kinase gene (PGK)²² (Fig. 4b). pERT87-25 was hybridized to a more

extended panel of human fetal tissue RNA samples and was found to recognize a broad size-range of RNA transcripts from small intestine and lung tissue (both of which have some smooth muscle) but not to RNA isolated from fetal brain, liver, spleen, kidney, thymus, heart and adrenal gland, nor to RNA from human transformed B and T cell-lines (data not shown).

In order to substantiate the RNA transcripts detected by pERT87-25, an oligo (dT) primed cDNA library was constructed²³ in the phage vector λ gt11²⁴ using the poly(A)⁺ RNA sample (Fig. 4a, lane 4). The resulting library of approximately 1.7×10^7 independent phage clones was amplified and 1.5×10^6 phage clones were screened²⁵ with both pERT87-25 and pERT87-4 as hybridization probes. Twenty phage positives were isolated, all of which hybridized only with pERT87-25, and of these twenty positives 5 cDNA clones with different sized inserts ranging from 1.0 to 2.9 kb were characterized further. The largest cDNA clone (2.9 kb) gave a repetitive pattern of hybridization to genomic DNA, and must be subcloned to find unique fragments which can be utilized on Southern blots. The remaining cDNA clones hybridized to human genomic DNA with common *Hind*III restriction fragments that map to Xp21; two hybridization examples are given in Fig. 5. Fig. 5a depicts the hybridization of cDNA number 5 (a 1.0 kb insert) that detects eight different restriction fragments mapping to the X chromosome which are all missing in the DNA isolated from B.B., the male patient with a visible Xp21 deletion and three X-linked diseases¹¹. The Southern blot in Fig. 5b shows the hybridization of cDNA 6 (a 2.1 kb insert) which detects the same eight *Hind* III fragments as cDNA 5 plus a ninth *Hind*III fragment of 4.2 kb that also maps to Xp21. cDNA 6 as well as two others (not shown) hybridize to several restriction fragments which do not map to the X chromosome. These autosomal hybridizing fragments may result from ligation artefacts in cDNA cloning or may represent homology of the Xp21 specific cDNA sequences to an autosomal locus. The nucleotide sequence of the cDNAs is currently being determined.

The cDNA clone 5 that detected eight Xp21-specific *Hind*III fragments in genomic DNA was then used as a hybridization probe (Fig. 4c) on the same Northern blot in which pERT87-25 identified a large transcript in fetal skeletal muscle RNA (Fig. 4a). The cDNA hybridizes to the same 16 kb transcript and smear underneath, more intensely than the small exon of pERT87-25, and with comparable intensity to that of PGK (Fig. 4b), after four days of autoradiography. The hybridization of pERT87-25 to the large transcript in Fig. 4a is at the lower limit of detection on Northern blot analysis.

To determine whether the cDNA clones represented transcription from regions of the DXS164 locus other than the conserved pERT87-25 segment, cDNA 5 was hybridized to phage digests representing the entire 220 kb of the DXS164 locus. The cDNA clone 5 detected the original pERT87-25 segment (a 3.0 kb *Hind*III fragment) as well as 6 additional restriction fragments

whose locations in the DXS164 map are shown at the bottom of Fig. 5. Examples of the hybridization of the cDNA clone to phage digests are given in Fig. 5c and d. In a similar analysis, cDNA 6 hybridized to an additional 4.2 kb *Hind*III fragment in phage DNA on the centromere side of the DXS164 locus which was not detected by cDNA 5. The potential exons of cDNA 5 are scattered over at least 110 kb and the exons of cDNA number 6 over 130 kb of the DXS164 locus with an average intron size of ~16 kb. These genomic distances do not include two Xp21 *Hind*III fragments seen on genomic Southern blot analysis (Fig. 5a and b) that could not be mapped to the phage walks from cloned DNA of the DXS164 locus. The spacing of potential exons in the DXS164 locus represents an order of magnitude difference between the size of the cDNA clones and the genomic DNA in which the individual exons lie. If this trend holds for the entire 16 kb RNA transcript, then the genomic locus for this gene could cover from 1,000 to 2,000 kb. A large size for the DMD locus is consistent with the high frequency of mutation seen in the disease²⁶, the heterogeneity of translocation breakpoints in Xp21 which give rise to DMD or BMD in females¹⁰, the 5% recombination of restriction fragment length polymorphisms (RFLPs) from the DXS164 locus with DMD and BMD mutations segregating in families^{15,27}, the large size of deletions detected in DMD males^{12,14,15} and the heterogeneity of deletion breakpoints¹⁵. The cDNA clones isolated with pERT87-25 are candidates for portions of the gene that is deleted, structurally altered or mutated to give rise to the phenotype of DMD and BMD. The most characteristic aspect of both disorders is progressive muscle degradation which may reflect expression of the DXS164 locus in fetal muscle. The cDNA clones are currently being used to search for similar transcripts in muscle isolated from human adult tissue and mouse tissues at various developmental stages. The use of small restriction fragments at the ends of the present cDNA clones to screen the same or newly constructed fetal skeletal muscle cDNA libraries, will allow the acquisition of the entire cDNA representation of the presumed large transcript. The nucleotide sequence and predicted protein product from this gene will help in the understanding of its natural function in muscle.

We thank Bruce Wentworth and Lydia Villa-Komaroff for the gift of the mouse recombinant phage library, Christine Disteché for a hamster-mouse somatic cell hybrid bearing a mouse X-chromosome, Fred Bieber and Robert Brown for fetal and adult tissue samples and Ulrich Müller and Stuart Orkin for a critical reading of the manuscript. This work was supported by The Muscular Dystrophy Association of America (L.M.K.), the NIH (NS23740) (L.M.K.), (HD18658) (L.M.K., D.M.K. and R.L.N.), (R01 20118) (D.M.K.), an RCDA to D.M.K., (HD20118) R.L.N. and the National Down Syndrome Society (R.L.N.). A.P.M. is supported by PHS NRSA (2T 32 GM07753-07) from the National Institute for General Medical Sciences and The Muscular Dystrophy Association of America.

Received 25 July; accepted 19 August 1986.

- Davies, K. E. *et al.* *Nucleic Acids Res.* **11**, 2303-2312 (1983).
- Aldridge, J. *et al.* *Am. J. hum. Genet.* **36**, 546-564 (1984).
- deMartinville, B. *et al.* *Am. J. hum. Genet.* **37**, 235-249 (1985).
- Bakker, E. *et al.* *Lancet* **ii**, 655-658 (1985).
- Kingston, H. M. *et al.* *Hum. Genet.* **67**, 6-17 (1984).
- Brown, C. S. *et al.* *Hum. Genet.* **71**, 62-74 (1985).
- Dorkins, H. *et al.* *Hum. Genet.* **71**, 103-107 (1985).
- Fadda, S. *et al.* *Hum. Genet.* **71**, 33-36 (1985).
- Goodfellow, P. *et al.* *Cytogenet. Cell Genet.* **40**, 296-352 (1985).
- Boyd, Y. & Buckle, V. J. *Clin. Genet.* **29**, 108-115 (1986).
- Francke, U. *et al.* *Am. J. hum. Genet.* **37**, 250-267 (1985).
- Ray, P. N. *et al.* *Nature* **318**, 672-675 (1985).
- Kunkel, L. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 4778-4782 (1985).
- Monaco, A. P. *et al.* *Nature* **316**, 842-845 (1985).
- Kunkel, L. M. *et al.* *Nature* **322**, 73-77 (1986).
- Frischauf, A. M. *et al.* *J. molec. Biol.* **170**, 827-842 (1983).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
- Ruskin, B. *et al.* *Cell* **38**, 317-331 (1984).

- Keller, E. B. & Noon, N. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7417-7420 (1984).
- Knott, T. J. *et al.* *Science* **230**, 37-43 (1985).
- Breitbart, R. E. *et al.* *Cell* **41**, 67-82 (1985).
- Michelson, A. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 472-476 (1983).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
- Young, R. A. & Davis, R. W. *Science* **222**, 778-782 (1983).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Moser, H. *Hum. Genet.* **66**, 17-40 (1984).
- Fischbeck, K. *et al.* *Lancet* **ii**, 104 (1986).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Bolivar, F. *et al.* *Gene* **2**, 95-113 (1977).
- Vierra, J. & Messing, J. *Gene* **19**, 259 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Bruns, G. A. P. *et al.* *Adv. exp. Med. Biol.* **154**, 60-72 (1982).
- Sanger, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Chirgwin, J. M. *et al.* *Biochemistry* **18**, 5294-5299 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Singer, M. F. *Cell* **28**, 433-434 (1982).

Archaeology gets graphic

from B.S. Ottaway, L. Sawyer and A. Miller

Interactive computer graphics designed for protein crystallography has also helped visualize a 5,000-year-old excavation site in three dimensions. Tomorrow's archaeologists stand to benefit from this marriage of the very old and the very new.

THE excavation of an archaeological site is a three-dimensional process that is usually visualized by two-dimensional section drawings and plans. A modern excavation can entail the meticulous three-dimensional recording of thousands of artifacts, together with elaborate environmental and stratigraphic data. With such recording, complete three-dimensional reconstruction of the stratigraphic evidence, which was destroyed in the course of an excavation, is possible in theory.

In practice, however, one of the major drawbacks in archaeology is that the three-dimensional relationships between individual stratigraphic layers, and between these layers and the artifacts retrieved from them, are difficult to study after excavation. This handicap has become much greater as excavations have moved away from stone-built structures towards the interpretation of large, spatially complex settlements.

Hence any method that conveys the three-dimensionality of an archaeological site would be a powerful tool in the final interpretation. Several centres are currently trying to devise a form of three-dimensional visualization based on graphics for micro- and mainframe computers. So far, the programs have mostly been used for pattern recognition and spatial analysis¹⁻³. None has used high-resolution graphics allowing real-time manipulation or interactive software designed for investigating three-dimensional problems. The IBM (UK) Scientific Centre at Winchester has worked along similar lines, but only with non-standard programs⁴.

Borrowing graphics

A program that has recently transformed the model-building stage of protein crystallography seems to have particular utility in archaeology as well. The program, called FRODO⁵, was written to display three-dimensional contour maps of electron density into which stick representations of a protein molecule can be built. Protein crystallography models have much in common with archaeological visualization, because data points relating to thousands of atoms must be fitted into a three-dimensional pattern that is not predetermined. FRODO uses depth-cueing, stereopsis and real-time manipulation to enhance the three-dimensional image, and colour to present several types of information simultaneously.

We conducted a feasibility study on using FRODO to represent archaeological data with the high-resolution interactive graphics facility in our department of biochemistry. For the investigation, we chose a small but complex archaeological feature of a prehistoric settlement site in Bavaria. The feature, a pit, was situated directly underneath the only entrance to a late Neolithic fortified enclosure. In accordance with standard excavation practice, the three-dimensional information about each layer and each artifact within the layers had been recorded⁶. Plans and sections had been drawn (Fig. 1) and an extensive photographic record built up during the excavation.

We converted the observed archaeological layers to three-digit codes denoting positions on a grid of points throughout the layers. FRODO was able to contour these data at up to seven levels (Fig. 2) and display each in a different colour. Artefacts were designated as water molecules, which was the simplest way of adding points using the unmodified software. Our system was an Evans and Sutherland PS300 interactive graphics system linked directly to a Digital Equipment Corporation VAX II750 computer running under the VMS operating system.

Timely retrieval

FRODO's images immediately disclosed several previously undetected patterns within the pit excavation. Colour-coding clarified the connectedness of a

number of layers which had, on excavation, appeared discrete. The patterns confirm that the pit was filled, partially re-cut and refilled during successive periods 5,000 years ago.

In addition to the stratigraphic advantages of colour-coding, FRODO's graphics also pointed out correlations between groups of artefacts and their matrices. With FRODO, it was easy to identify artefacts located in inappropriate layers. The arrangement of artefacts could also be quantified in calculations of densities and spatial distribution patterns. And with real-time manipulation, FRODO allowed viewing of individual layers or segments of layers, and could zoom in on or rotate the display.

The most useful aspect of FRODO, however, is that representations can be stored on magnetic tape and quickly recreated when new information becomes available. For example, preliminary results suggested that the pit had a ritual function, but more detailed analysis of the bones discounted this and led to a reassessment of the dome-shaped layers. When pottery analysis is complete, it will also undoubtedly alter the interpretation.

This preliminary study indicates that many archaeological questions can now be efficiently investigated after excavation by studying and manipulating the three-dimensional aspects of the excavated feature in a way that is not normally possible owing to the destructive nature of the excavation itself. Our approach needs to be

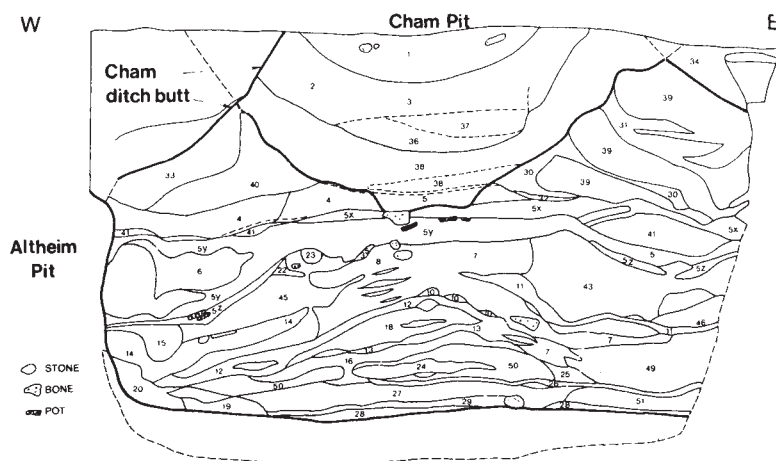


Fig. 1 Conventional archaeological section drawing of a pit within a prehistoric site in Bavaria. Two Neolithic groups, the Altheim and the Cham, were occupying the site. The section drawing shows the dome-shaped earliest deposits (the 'Altheim pit'), subsequent re-cutting and re-filling of the pit by the first occupants of the Cham group (the 'Cham pit') and final truncation by the latest occupants of the Cham group (the 'Cham ditch').

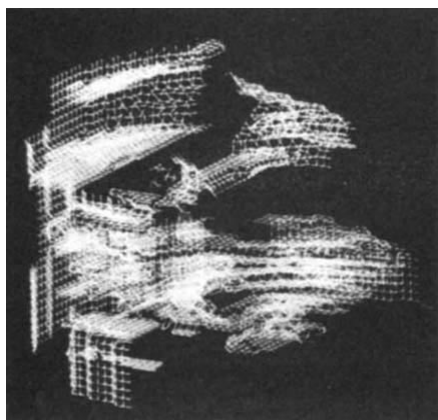


Fig. 2 FRODO images of seven of the archaeological layers shown in Fig. 1. The dome-shaped nature of the earliest, lowest layer, the re-cut of the Cham pit and the final truncation caused by the construction of the Cham ditch (in this figure in reverse image) are shown here in black and white. In reality, when working with FRODO, these seven layers are shown in colour on the screen.

developed, most notably with regards to sampling intervals. We arbitrarily chose a grid of $5 \times 5 \times 5$ cm, but the interval need not be isotropic.

We used FRODO because FRODO was available. Obviously, a program dedicated to archaeological applications could yield even more satisfying results. □

Barbara Ottaway is at the department of Archaeological Sciences, University of Bradford, Bradford, West Yorkshire BD7 1D8, UK. Lindsay Sawyer and Andrew Miller are at the Department of Biochemistry, University of Edinburgh, Hugh Robson Building, George Square, Edinburgh EH8 9XD, UK.

1. Hodder, I. & Orton, C. *Spatial Analysis in Archaeology*, (Cambridge University Press, Cambridge, 1976).
2. Kampffmeyer, U. *Saalburg-Jahrbuch* **40-41**, 121-133 (1984/85).
3. Laffin, S. & Sutton, N.D. *Proc. Computer Applications in Archaeology*, 111-113 (University of Bradford, 1983).
4. Colley, S. & Todd, S. *Archaeological Computing Newsletter* **4**, 3-8 (1985).
5. Jones, T.A. *J. Appl. Cryst.* **11**, 268-272 (1978).
6. Ottaway, B.S. *Archaeologisches Korrespondenzblatt* **14**, 23-29 (1984).

Tokyo technology

At the end of October, over 300 companies will convene in Tokyo for the 21st scientific instrument show. Below are some of the top exhibits.

Shimadzu has a pair of **gas chromatographs** geared towards laboratory automation (Reader Service No. 100). The GC-15A/16A series can operate independently or as core instruments in multi-GC and computerized systems. With a 9-inch

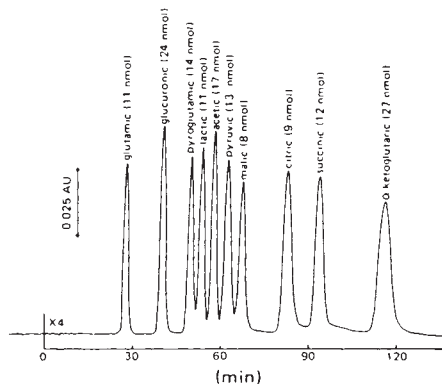


colour CRT, a menu-driven dialogue between operator and machine can establish chromatographic protocols. Shimadzu makes sample injectors in modular designs so they can be interchanged; up to four detectors can be installed on either model. Both models feature up to four analogue outputs and self-diagnostic capabilities.

For automatic counting and classifying colonies of microorganisms, Toyo engineered a video scanning **image analyser** that can also transplant designated colonies to other areas of the dish (Reader Service No. 101). Toyo's model CT-400 selects colonies for transplantation on the basis of the density, shape and size that the user specifies. The inoculation plunger is automatically heated and cooled after each transfer to maintain sterile conditions. Starting with a stack of as many as 500 Petri dishes, the analyser moves each dish to the scanning position, removes the cover and performs the transfer with no external assistance.

Mediclave **steam pressure sterilizers** have three programmable cycles for different modes of autoclaving (Reader Service No. 102). The HK-II-E series models, with one or two doors and a built-in generator, is cabinet-standing but can be wall-recessed as well. Hirayama Manufacturing, which makes the Mediclave line, says the chamber is of double-wall pressure shell construction, covered with fibreglass to reduce heat loss. Some of the models have removable stainless steel shelves, for which Hirayama can also provide a transfer wagon.

A little bit of everything for the laboratory makes up Tokyo Rikakikai's new **catalogue** (Reader Service No. 103). The company manages to detail its evaporators, chromatographs, fermentors and freeze driers in 30 pages. Most of the instruments and equipment are shown in full colour, with accompanying diagrams and



illustrations of machine performance. Rikakikai also defines product specifications and accessories where available.

Fluoro Plastic Devices for Trace Analysis

Berghof

Teflon® PTFE, PFA and FEP

excellent materials for non-contaminational working with pure and aggressive substances at high temperature

- „Subboiling“ Acid Purifying also for hydro fluoric acid
- Storage vessels for liquids (5 to 45 litres)
- Decomposition Systems under pressure, reflux conditions
- Separation of volatile compounds up to 260°C
- Sample preparation (phase separators, gas sample valve)

Berghof GmbH · Abt. Labortechnik

D-7412 Eningen · Harretstr. 1 · Tel. 07121/89010 · Telex 729529



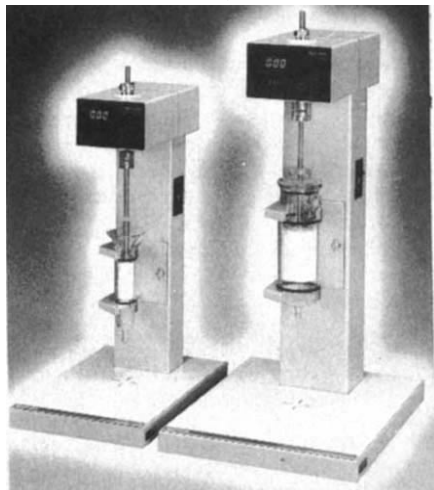


Sun cell in the jungle?

PORTABLE microcomputers powered by photovoltaic cells are aiding efforts to save tropical rain forests in Papua New Guinea (Reader Service No. 104). Solar panels like the one pictured above cost roughly £80 (UK) and allow daily use of the computers without pauses for recharging. Using panels supplied by Chronar Ltd, the natives of the rainforest region of Daru have been able to record and transmit data on indigenous plants and cropping techniques that could vanish through deforestation. Their work is supervised by Operation Raleigh, a British organization that hopes the data will underscore the value of biological diversity.

In a whirl

A combination of design features helps Yamato's low shear homogenizers disrupt cells without disrupting organelles and other cell structures (Reader Service No.

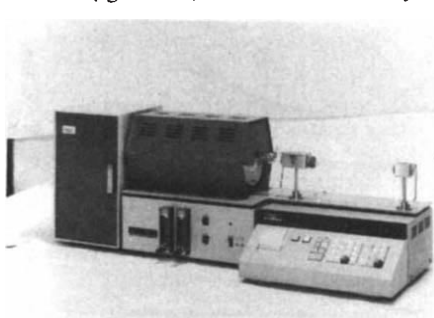


105). The company's continuous flow models LH-21 and the larger LH-41 use only one passage through tissue to minimize the heat and shear to which cells are exposed. Yamato says the grooves on the Teflon pestle, which can achieve speeds of 1,200 r.p.m. or faster, force tissue through the shear region of the pestle. Digital speed control and a clamping configuration that replaces manual holding contribute to reproducible results. Yamato also has cylinders designed for continuous injection of bulk tissue samples.

Kubota Corp makes a refrigerated centrifuge with 20,000 r.p.m. maximum speed and automatic r.p.m. calculation from specified g-force (Reader Service No. 106). There are nine user-programmable channels built into Kubota's KR-2000C, and samples stay at 0°C at any speed. Kubota says the centrifuge contains an integrator system and self-diagnostics via its microprocessor control system. The g-force calculation works at less than maximum radius as well.

Automating concentration

A microprocessor controlled coulometer from Mitsubishi Chemical Industries will measure total chlorine or total sulphur content in liquids or solids (Reader Service No. 107). The TSX-10, with titration speed maximums of 73 µg Cl/min and 66 µg S/min, has three electrolytic



current levels giving a sensitivity range from p.p.m. to per cent. Measurement takes from 5 to 20 minutes, according to MCI, and the internal microprocessor gives calculations and print-outs of concentrations. The TSX-10 works by the method of oxidative pyrolysis and coulometric titration. Its measuring range spans 0.05 to 300 µg for S and 0.1 to 500 µg for Cl. MCI equipped its instrument with digital display of results in addition to print-outs.

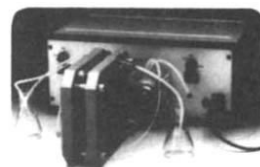
A digital concentration meter that takes measurements by light refraction will also determine mixing ratios, reaction end-points and concentration variations (Reader Service No. 108). Atago Ltd's PR-1, which weighs 285g, contains a microcomputer that automatically compensates for temperatures between 10°C and

35°C. Fresh water calibrates the instrument, and measuring time, according to Atago, is approximately 3 s. The company says its refractometer is ideal for monitoring concentration changes due to evaporation, volatilization or degradation; will assist in inspecting concentration changes in culture solutions and pinpointing terminal concentrations in fermentation; and can verify dilutions. PT-1 minimum sampling volume is 0.1 ml.

ADVERTISEMENTS

LIPOSOMES

REPRODUCIBLE - HOMOGENEOUS - EXTREMELY RAPID



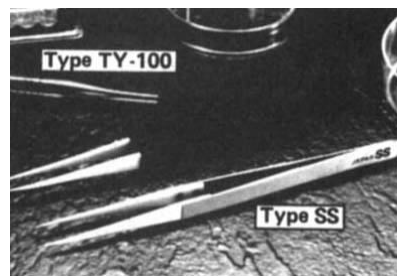
Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.26

NEW CERAMICS PRECISION: FUNA-TWEEZERS



"FUNA-TWEEZERS" can eliminate all demerits of stainless or plastic tweezers. The stainless stems of Type SS and the Derlin stems with glass reinforcing agent of Type TY-100 give appropriate elasticity.

These show excellent features such as high anti-corrosiveness, heat insulation, high flame resistance, etc. These are also free from plastic deformation, metallic contamination and magnetization, etc.

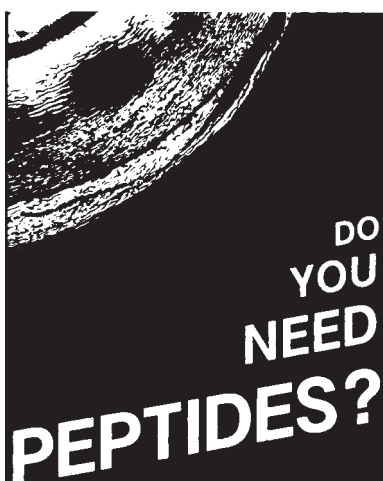
These tweezers are contributing to the industries and research such as Biotechnology, Electronics, etc.

FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone : Tokyo 03-293-2352
Telefax : 81-3-295-5545
Telex : J28489 FUNA

Reader Service No.12

ADVERTISEMENTS

Roland GRUNBERG - NANCY - FRANCE



DO
YOU
NEED
PEPTIDES?

NEOSYSTEM
LABORATORIES

PROVIDE
A COMPLETE SERVICE FOR
PEPTIDE SYNTHESIS
PEPTIDE AND PROTEIN SEQUENCING
PEPTIDE ANTIBODY PRODUCTION

WE PROVIDE ALSO
CONTRACT RESEARCH SERVICES
TO INDUSTRIAL AND ACADEMIC
INSTITUTIONS



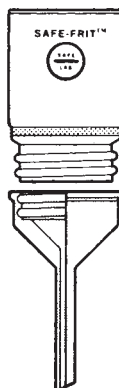
NEOSYSTEM
LABORATORIES

B.P. n° 309 R8
67008 STRASBOURG Cedex
FRANCE - Télex 880 088 F
Tel.: 88.36.96.96

Reader Service No.4

SAFE-FRIT™ FUNNELS

Safe-Lab's NEW fritted glass funnels feature a removable, unbreakable polypropylene stem just below the fritted glass disc. This unique design eliminates the danger of broken glass funnel tips and permits easy cleaning from both sides. Removable stem allows you to set the funnel top on counter or directly inside an oven or desiccator for drying, without the danger of tipping or spilling. Polypropylene has excellent chemical resistance and threads will not leak, even at high-vacuum.



Call or write for a FREE catalog:
Safe-Lab, Inc., P.O. Box 1290,
Santee, CA 92071
(619)448-9100

Reader Service No.28

SINGER INSTRUMENTS
MICROMANIPULATORS & ACCESSORIES
YEAST?
TETRAD ANALYSIS?



Singer MK III Micromanipulator
and Precision Elevator.

Write direct to:
Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

Versatile Tissue Culture
Incubator

Modular Incubator
Chamber (the Cell/
Sphere™) for all
tissue cultures,
aerobic or
anaerobic.

Low
Cost



Eliminates evaporation loss and contamination. Easy to use - flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **biFlups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.**

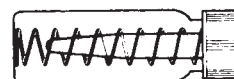
Reader Service No.7

British wares

Chromacol Ltd, a first-time visitor in Japan, has developed **tapered vials** for both routine and microsampling work (Reader Service No. 109). Chromacol thinks these vials will save time and money by standardizing to one vial for both kinds of work. The vials range from 6 ml down to 200 µl and can be used with over 50 kinds of autosamplers, according to the company. Chromacol will also supply seals for different detectors, solutions, storage applications and so on. Most of the vials are glass, although screw-top and crimped-top polypropylene varieties are also available. Polypropylene screw-tops for the glass vials have recently joined Chromacol's range as well.

Automated thin layer chromatography analysis is just one of Laboratory Impex's offerings (Reader Service No. 110). In addition to the LB 285 automatic TLC machine, several other instruments have come out of the company's research group recently, such as an international standard contamination monitor and an HPLC radioactivity flow monitor. Tecan robotic sample processors now come in a size for small batches, says Lab Impex, which also markets monoclonal antibodies and special reagents such as fluorescence-activated cell sorters and analysers. Finally, the company has recently developed two non-radiometric automated blood culture systems coupled with DEC Pro 380-based data management capabilities.

Phase Separations Ltd will bring its latest **process-scale LC packing material** to Tokyo (Reader Service No. 111). The company says its silica particles give high



efficiency and low back pressure and are available as plain silica or with C₁, C₆, or C₁₈ bonded phases. Spherisorb VLS has a particle diameter of 20 µm, low metal ion content and a pore diameter of approximately 300 Å. In view of this combination of traits, Phase Separations recommends the material for protein separations. The company also distributes a guide to packing with Spherisorb materials. Other Phase Sep products, including a variety of autosampler vials, will be at the Tokyo show as well. □

These notes are compiled by Karen Wright from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.